

Life with the new technologies

Once the issues have been clarified and irrational protests brushed aside, the rules by which to regulate the transfer or experimental investigation of human embryos almost write themselves.

THE British Government is but one of several that have set up committees to examine the ethical problems of new techniques in biology. From the State of Victoria in Australia to France and the United States, groups of careworn people chosen for their disparateness are hard at work on schemes for regulating the use in people of techniques such as *in vitro* fertilization and embryo transfer. So it is not in the least surprising that the Warnock committee, set up a year ago to advise the British Government, should apparently not yet have agreed what it will say when it reports in June. The snag, for the committee and the government, is that the process of winning agreement by waiting for strong-minded people to batter each other into consensus is slower than the development of the technology whose regulation is intended.

A few simplifying principles may help. The sources of anxiety should be identified and recognized, if only because no public policy can command respect and thus assent unless it accommodates all reasonable opinions either by concession or by reasonable argument. Religious objections are the most evident; Catholics have specific difficulties about the definition of human life and vigorous reservations about its destruction in any form (even as an eight-cell embryo), but many other religious groups strongly hold that only natural methods of procreation are permissible. Lawyers have other difficulties. Whose is the child grown in one woman's uterus from another's oocyte? Physicians of a eugenic cast of mind worry about the risks of inbreeding from the over-use of sperm banks. Others complain that to the extent that some people may profit from the new technologies, society will be debased. Still others hold, irrationally, that any technique with its origins in a laboratory must be an abomination. Not all of these opinions can be reconciled, but each of them must be met.

Deprivation

The Warnock committee and the others will not succeed unless they provide a clear statement of the reasons why the technology on which they are brooding has become a pressing issue. For most of the developments now in prospect, human infertility and congenital diseases are the driving forces, and should not be underestimated. Can it make sense that relief from childlessness or the avoidance of genetic defects should be denied? The benefit (to parents in particular, but the rest of us as well) that might flow from these developments cannot easily be overestimated. Yet for a long time to come, further innovations are unlikely to take precedence over a technique now well established — diagnostic amniocentesis followed, where necessary, by prophylactic

abortion. And while there are many sections of society that would not consider using these techniques themselves, there seems to be no great antagonism between the opposing sections of the communities in the countries in which they are used.

Animal experiments

The scientific need for access to human embryos should be acknowledged to be less urgent. Animal experiments will for the time being serve most scientific needs. So little is as yet known of the mechanism of differentiation that ambitions to learn how to optimize the specifically human process of intrauterine development cannot now be matched by pointed programmes of observation and experiment. Yet that time cannot be long delayed. And the benefits could be greater than the eugenicists have ever (falsely) promised.

So what should a beleaguered committee recommend? The first need is to insist that the cant and inconsistency hidden in some of the objections to the use of the new technology should be brought into the open. Thus complaints that the new developments will encourage women who choose to live without men to produce children outside a nuclear family, which is already possible with the assistance of a sperm bank, are thoroughly inconsistent in societies that prolifically generate "single-parent families" by natural means and are inequitable as well. Complaints that the new techniques are objectionable because some may earn a living from them, and the implication that they would be more acceptable if no money changed hands, are frankly hypocritical. People's time, trouble and skill are required. To insist that they should be offered free would be to ensure that they were spent on other tasks.

Within such a framework, rules for the regulation of the new technology almost write themselves. Physicians and their professional organizations have a central role to play. The principles, in a form general enough to apply to foreseeable developments as well as those now used (amniocentesis included) should be these.

Because public anxiety about the new developments is real, the avoidance of frivolous applications of techniques must be the first objective. This is the basis on which most of the legislation for the regulation of abortion is already based. Some persuasive reason is required for each legal abortion, with the beneficial social result that such arguments as persist centre on the precise conditions that should apply. With the new techniques, exactly the same principle should be followed. Thus women demanding embryo transfer to a surrogate mother for the sake of avoiding the risk or even the inconvenience of pregnancy should be denied. Whether the predetermination of the sex of a child should be counted as frivolous is not easily decided; in any case, the issue is probably academic if only because the separation of sperms that determine male or female sex must soon be possible.

If physicians must shoulder the day-to-day responsibility for the avoidance of frivolity, professional associations and licensing authorities must in future function more vigorously as monitors of what physicians do than is at present the case. Although physicians set great store by the privacy of the relationship between a physician and his patient (witness the just battle of the British Medical Association with the British Government over the misnamed data protection bill now in the House of Commons), there is a strong case that the nature and number of cases of

Nature in Tokyo

Nature has opened an office in Tokyo. Dr Alun Anderson, for the past few years in charge of the News and Views section of *Nature*, will be in charge of the office for an initial period of one year. By this means, *Nature* hopes more efficiently to cover news of developments in Japan and also to provide tangible help to Japanese scientists wishing to submit research papers or other communications to *Nature*.

Dr Anderson's address is *Nature*, Macmillan Shuppan KK, Eikow Building, 10-9, Hongo 1-chome, Bunkyo-ku, Tokyo 113 (Telephone (03) 816 3756/7; Telex JATOK J32121 MSKK). □



unnatural embryogenesis dealt with by physicians should be open to scrutiny within the medical profession. The usual procedure by which a physician engaged on a novel treatment protects against criticism — the second opinion — will probably not suffice in these novel circumstances. There is also a need for formal arrangements whereby cases can be discussed, necessarily after the event, within a wider professional circle.

Anonymity

Even the traditional anonymity of the patient may not be inviolable, but for legal reasons. The question arises from the operation of sperm banks like those in the United States but also from the more common if more clandestine use of AID (artificial insemination by donor). The abstract legal question whether the offspring of such a procedure may have a right to claim the paternity of the donor of the semen should be settled by straightforward legislation (laying it down that no claim can lie). The eugenicists' fear that the result may be a major inbreeding is sustained by the likelihood that AID donors may be used repeatedly. But there is no reason to think that even AID is now practised on a scale that would invite these dangers. The more substantial reason for knowing what sperms have fertilized what oocytes is that with the inevitable improvement there will be of people's understanding of human genetics, the need may arise to be able to reconstruct the genetic history of people not yet born or even conceived. It may be prudent to identify donors by number if not by name, which implies some mechanism for keeping records more substantial than physicians can provide.

The problem of the surrogate mother is said to be one of the particular difficulties with which the Warnock committee is wrestling. As things are, the problem is not numerical but one of principle. Suggestions that the practice should simply be made illegal are probably inequitable and certainly impractical. Why should women infertile because of some uterine defect be denied the chance of producing a genetically authentic child with the help of some other uterus while women whose fallopian tubes are blocked have access to *in vitro* fertilization? And while neither of these techniques is technically simple or free from risk, there is every chance that they will in due course be internationally accessible. Legislation is neither wise nor enforceable.

Would-be users of a third party's uterus have quite a different worry — the fear that an embryo may be conceived, grown to term and then not be handed over to the genetic parents, whatever the terms of the contract, explicit or otherwise, made at the outset. Even this problem is not, however, as unfamiliar as it seems. One of the most familiar if poignant tales told by the adoption agencies is of women resolved, when pregnant, that their offspring will be adopted who change their minds once the child has been born. At that point, genetic ties matter less than instincts. From this it follows that a woman who carries another's embryo should not be required by some contract made in advance to hand over the child that results from it, nor to be sued for breach of contract.

Self-discipline

The general rule for the regulation of unnatural embryogenesis by techniques now available or likely to become so is therefore straightforward. Tighter self-discipline within the medical profession coupled with a tendency towards explicitness for the sake of seemliness and so as to gather data are the necessary trends that should require only minor legislation to make them workable. For the plain truth is that while the technology of artificial embryogenesis is likely continually to become simpler, it will never be so simple as to be preferable to the natural method. The problem will never be numerically huge. Since the issues of principle that are raised by foreseeable techniques are not nearly as novel as they are made out to be, the committees brooding on the subject should above all seek not to exaggerate the importance of what they are about.

The regulation of research in this field is similarly a problem that can too easily be exaggerated. Laboratories with people working in the field of artificial embryogenesis should equip themselves with ethics committees for the scrutiny of experi-

mental protocols, and should prudently set these up off their own bats, not wait to be instructed to do so. One obvious criterion that should be satisfied by proposed experiments with human sperm and ova is that there should be good reasons why the same findings cannot be obtained with other mammalian material. Similarly, experimental protocols should demonstrate convincingly the reasons why embryos should be kept alive for whatever length of time is specified. For even though there is no logical accommodation to be found between the views of those who hold that even a fertilized ovum is a living creature and those who consider that observations of such systems are permissible, the practical truth is that the anxiety (and opposition) of those who hold to the extreme definition is bound to be moderated if researchers can exercise self-restraint. □

Smoke-screen tactics

An advertising campaign in support of smoking is best challenged with detailed facts.

THE tobacco industry has what is known in the advertising industry as an image problem. Despite valiant efforts to counter the perception that cigarette smoking is the single largest preventable cause of cancer (even for cowboys on horseback), the industry's production of cigarettes in the United States fell by 7 per cent in 1982. The percentage of adult Americans who smoke fell from 39 per cent in 1982 to 29 per cent just a year later; overall, 32 million Americans who once smoked no longer do so.

These figures have the industry worried and have spawned a new advertising campaign from R.H. Reynolds, one of the largest tobacco companies in the United States, aimed at its best market: smokers who are hooked, who are ready to quit under the weight of health evidence and who are desperately eager for some straws to clutch at. The first advertisement in a planned series calls for an "open debate" on smoking and health, hinting at scientific evidence that legitimately questions the establishment view of the hazards of smoking and quite brilliantly sows the impression that the subject is still a controversial one. The advertisement, which is short on facts, is bound to give those smokers who are looking for an excuse not to quit something to grasp at, namely the notion that smoking and health is just another of those issues upon which scientists can never agree. The tactic is simple and the damage is done in the first instance.

The American Cancer Society, the American Heart Association and the American Lung Association are doing their best to counter the campaign. They have likened the industry's call for an "open debate" on smoking to a call for open debates on the lethality of bullets or on whether the Second World War took place. Predictably, the industry has cried that its rights of free speech under the first amendment to the US constitution are being trodden upon by the health organizations — thus adding to the air of controversy it is hoping to generate and to its cultivated image as an oppressed minority whose legitimate views are being squelched by a dogmatic establishment.

The only hope of countering the tobacco company's cynical propaganda is the tireless repetition of facts, such as those found in the American Cancer Society's exhaustive Cancer Prevention Study, which tracked a million people for twelve years and which showed that no fewer than 83 per cent of lung cancers are associated with smoking. Even simple facts like these are hard for the industry to deal with. The standard lame response that epidemiological studies offer merely "statistical" evidence, not causal proof, of the danger of smoking should be provoked as often as possible — the more the industry must deal with details, the more it looks ridiculous and the less it looks responsible despite its attempts to do so by means of other advertisements, warning the public that smoking in bed is dangerous and that smoking is an "adult" pursuit. Only when the tobacco industries acknowledge that smoking, in bed or out, is dangerous and cease to advertise it as a pursuit of cowboys, the young, the athletic and the attractive, will they become truly responsible. □

Satellite launchers

Japanese plan to join the big space league by 2001

Tokyo

PLANS to put Japan alongside other nations with a major investment in space were heralded by an ambitious 15-year development programme approved by the Space Activities Commission last week. The programme should turn Japan into a major space nation by the end of the century and enable it to compete with the European Ariane and the US space shuttle for the launch of large communications satellites early in the next decade.

At the heart of the programme is the decision to develop a new rocket — the H2 — powered by a domestically developed cryogenic oxygen–hydrogen engine, capable of putting a 2,000-kg satellite into geosynchronous orbit. Development costs are estimated at 200,000 million yen (\$870 million) and will necessitate big increases in the National Space Development Agency (NASDA)'s annual budget of ¥130,000 million.

In addition, the programme calls for participation in building the US space station — although to what extent is not yet clear. Earlier Japanese plans for a manned space flight are discarded as uneconomic but provision is made for continuing use of the space shuttle for development of new materials under zero-gravity conditions. The new programme — the first to be announced for six years — was given the go-ahead at a meeting of the Space Activities Commission chaired by Michiyuki Isurugi,

director general of the Science and Technology Agency, which is responsible for all of Japan's "big science".

The decision to build the H2 rocket stems from confidence that Japan has mastered much of what it can learn from the United States in the field of launcher development. After a late entry into the space race in 1969, NASDA rapidly acquired the capability to launch small satellites by the aggressive import of US rocket technology. Its first launch vehicle, the three-stage N1, was essentially a copy of the Thor-Delta and could lift only a 130-kg payload into geosynchronous orbit. The N2, which replaced it in 1981, brought payload capacity up to a modest 350 kg. But while the N series rockets were being built under US licence, the Japanese were developing the H1, a far more powerful cryogenic oxygen–hydrogen engine that would mix Japanese and US technology and be capable of taking a 550-kg payload into orbit by the mid-1980s. Next in line would have been an extended version, the H1B, designed to take an 800-kg payload. But rapid development of the cryogenic engine has allowed the bold decision to skip the H1B in favour of the new H2, to be ready soon after 1990.

Success will give Japan the capacity to launch 2,000-kg satellites — the same capacity as the Ariane 4 rocket planned for completion around the same time and only a little less than the 2,270-kg capacity of the

Illmensee inquiry

CORRECTION

In the article under this heading on page 673 of last week's *Nature* (23 February), a word was inadvertently omitted from the second sentence. It should have read: 'But in a report that is highly critical of Illmensee's experimental records, the six members of the commission are unable to reach agreement on whether "the hypothesis that the protocols were fabricated could be upheld"; some felt it could not, others considered "there was no compelling evidence to support or refute the hypothesis"'. The word "not" was omitted due to an error in the editorial process. We apologize for the error and any distress it may have caused. □

space shuttle. And by using domestic technology, Japan will be free from legal restrictions on competing with the United States that come with building under US licence.

The entry of Japan into the large satellite launching business is bound to make both Europe and the United States nervous. There is also nervousness within Japan that the project is too great a gamble and that there will be insufficient business for all three launcher groups. At present, only large communications satellites reach the 2,000-kg mark and critics of the programme feel Japan would be safer sticking to the medium-sized satellite market. But the Japanese Government has already been irritated by US attempts to limit the transfer of satellite technology to Japan and independence from the United States is an important goal.

While the start of the H2 project now seems certain, the extent of Japanese participation in the US space station project is less clear. Hard bargaining is likely when James Beggs, administrator of the US National Aeronautics and Space Administration (NASA), arrives in Japan on 11 March to discuss the project. The Americans have reportedly asked Japan to take a 10 per cent share in the space station's development costs. Although the Japanese response has so far been encouraging, it seems unlikely with the present budget deficit that the Diet will approve expenditure on this scale — or that the Science and Technology Agency will tolerate the cuts in other projects that would probably become necessary.

The Japanese Government will also find itself in a serious dilemma if the US Government cannot guarantee that the space station will be used only for peaceful purposes. The Japanese constitution specifically forbids national military forces and the Japanese public is remarkably sensitive to any increase in military involvement with the United States. Some difficult negotiations over how much say the Japanese have in the deployment of the station are almost certain to lie ahead.

Alun Anderson

Japan versus Europe — size the key

JUST before Japan announced it would develop a big cryogenic launcher (H2 — see above), a senior French rocket engineer discussed the plan with the Japanese space agency, NASDA. The engineer, M. Jacques Borromée, is in charge of developing the cryogenic engine for the European Ariane 4 launcher — with which Japan believes H2 is in direct competition.

French reports described the meetings as "fruitful", but according to a spokesman for the French space agency CNES, the object of the visit was merely "information". The meeting was to be seen in the context of many meetings between CNES and other space agencies, he said, and nothing was said about possible commercial collaboration between CNES and NASDA.

However another French space engineer described the potential H2 payload of 2,000 kg in geostationary orbit as modest. Ariane 1 can achieve 1,825 kg, and Ariane 3 (to be first launched this summer) 2,580 kg. Ariane 4 with its cryogenic engines and four liquid boosters could put up a

4,200-kg communications satellite, the size adopted by the Intelsat corporation (a 108-nation body which effectively sets world standards for such satellites) for its next (Intelsat VI) series. Also, the first Ariane 4 proving flight is planned for 14 March 1986, four years ahead of any likely launch of the Japanese H2. This flight of Ariane 4 will be in the "most difficult" configuration, using two liquid fuel and two solid fuel boosters, with a potential 3,700-kg payload. If successful it will prove the full capability of Ariane 4.

The question becomes one of the weight of communications satellites by the end of the decade. In Japan there is a view that it will fall to around 2,000 kg, in part because of Japanese efforts to microminiaturize satellites. If successful, Ariane 4's maximum payload may become unnecessary but already it is planned to have members of the Ariane 4 family suited to a smaller payload by variations in the number and type of boosters carried by the first stage.

Robert Walgate

US research and development

Fears of failing competitiveness

Washington

THE robust increase in federal spending on research and development proposed in the Reagan Administration's 1985 budget request appears to have done little to cure the United States of its chronic fear of Japanese and European competition in high-technology industries. At a major conference on technology policy held here last week by the Institute of Electrical and Electronic Engineers, the air was thick with grim talk of the United States' slipping lead.

Dr George Keyworth, the presidential science adviser, did his best to raise spirits by pointing out that science had fared better than it had dared expect under the Reagan Administration. Over the four years of the administration, federal research and development spending will have increased by 52 per cent, he said. Basic research had grown by 55 per cent to a new high of nearly £8,000 million.

However, a much less upbeat message was delivered by Dr Simon Ramo, a co-founder and director of the TRW Corporation, a leading US high-technology company. The United States, he insisted, was losing its technological leadership. Proof was arriving "daily" from numerous fronts — decaying equipment in engineering schools, weaknesses in mathematics and science education and an "adversary" relationship between government and industry. Without vigorous government action, Ramo warned, US industry would lose its "technology war" with the Japanese. Important first steps would include relaxing the anti-trust laws, decreasing capital gains taxes and investing heavily in engineering education.

Coincidentally, the National Science Foundation (NSF) has just released a considerably more sober assessment of the United States' scientific and technological capabilities, in its annual *Science Indicators* report. The report concludes that, no matter how it is measured, the US lead over the rest of the world is still formidable.

In 1979, for example, the United States spent about as much on research and development as France, West Germany, Japan and the United Kingdom combined. It devotes a higher proportion of its gross national product to research and development and employs more scientists and engineers in this field than any other country except the Soviet Union.

One consequence of this investment is that US scientific papers dominate the contents of learned journals. In 1980 US authors wrote more than a third of the papers in scientific and technical journals. And despite a decade-long decline in patenting (a useful indicator of inventive activity), the US record remains

impressive.

So is the much-ballyhooed decline in US scientific leadership a mirage? Not entirely. For one thing, the huge US investment in research and development is heavily skewed in favour of defence. Defence already swallows up well over 60 per cent of federal spending on research and development, and that proportion is expected to increase. Japan and West Germany spend only a tiny fraction of their science budgets on defence research and development, which puts the ratio of their civilian research and development to gross national product ahead of the United States.

Nor is the US grip on scientific publishing as firm as it seems. The NSF report analyses more than 2,000 highly cited journals and finds that in six out of

eight fields the US contribution declined between 1973 and 1980. Only in biomedicine was there an appreciable increase.

The most disturbing decline, however, is in the amount of patenting by US inventors. Between 1971 and 1982, NSF reports, US domestic patenting decreased by about 40 per cent, while foreign patenting in the United States increased.

But the United States continues to provide assistance to the scientific endeavours of other countries. In the 1981–82 academic year, more than 150,000 foreign students were enrolled in scientific and technical courses in US universities.

It is not a one-way street, however. NSF says there is evidence that US scientists are gaining more from foreign research. In 1980, 44 per cent of all citations found in US publications were from foreign journals. That is an increase since 1973.

Peter David

Nuclear safety

Risky business of assessment

Washington

THE Nuclear Regulatory Commission (NRC) is moving cautiously towards expanding the use of the controversial technique of probabilistic risk assessment (PRA) to evaluate the safety of nuclear power plants.

The technique, which relates the known failure rate of nuclear components to the overall risk of such serious accidents as a core melt, has been sharply criticized for failing adequately to account for human error and for the intrinsic impossibility of imagining every possible route that could lead to a serious accident. The technique, which is widely used in industries such as aerospace, was first applied to nuclear plants in the 1975 Reactor Safety Study, directed by Norman Rasmussen of Massachusetts Institute of Technology; that report unleashed a storm of controversy, with critics calling it an empty numerical exercise whose only purpose was to lend an air of objectivity to foregone conclusions on the safety of nuclear power.

NRC is expected shortly to require all new plants to carry out a probabilistic risk assessment as part of the licensing process. Initially, these assessments would be used only as a guide to the design of safety systems and to help the plant's operators and federal inspectors to focus on the most vulnerable links in the systems. But within the next two years, NRC may begin to require probabilistic risk assessments as proof that plants comply — albeit within a generous margin of error — with NRC's previously adopted general safety goals for nuclear plants. These goals, which were adopted last year against the recommendation of NRC's own technical advisory committee on reactor safeguards, require that core melt accidents should not occur more often than once per every thousand

reactor-years of operation and that nuclear plants should add less than one-tenth of one per cent to the background risk of accidental death and cancer from all other causes.

A draft report commissioned by NRC* raises some doubts about the commission's strategy, however. While noting the usefulness of probabilistic risk assessments in identifying safety problems, the report warns against attaching much credence to their numerical results. It notes that uncertainties, which cannot even be quantified, make it "very difficult to determine formally whether a specific safety limit (in terms of public risk or frequency of a core melt) is met". And it warns that the sort of safety goals adopted by NRC inevitably causes disproportionate attention to fall on the "bottom line" of a risk assessment — the risk of core melt — despite repeated warnings about uncertainties.

Robert Bernero of NRC's risk analysis division says the commission is well aware of these limitations and that "no one is even suggesting the replacement of current regulations with PRA". The current safety regulations consist of engineering standards that specify the construction of components, for example. But Bernero says there is a possibility that NRC may adopt a requirement that reactors meet NRC's safety goals as calculated by probabilistic risk assessment, perhaps within an order of magnitude to allow for the inherent uncertainties.

After public comments are received on the report, it will receive peer-review from the National Science Foundation.

Stephen Budiansky

* Probabilistic Risk Assessment (PRA): Status Report and Guidance for Regulatory Action, US Nuclear Regulatory Commission, NUREG-1050 draft report.

Computer link-up

IBM fills the European gap

IBM, the American computer giant, is prepared to spend some millions of dollars to set up a computer network in Europe, linking universities and research centres. The similar US BITNET will also be linked to the European network, through nodes in Rome and at the City University of New York (CUNY). In this way, the establishment of the system will put researchers with computer access in communication from the west coast of the United States to Haifa in Israel (in the European net) and all points between.

According to IBM Europe, based in Paris, IBM will provide all the hardware and most of the software necessary to enable computer users in Europe to send files and electronic mail to each other. The network is being called "EARN" for the European Academic and Research Network, and will not be restricted to IBM users.

IBM will itself gain by being on the network — and thus in touch with a vast community of researchers — and by establishing greater "visibility" in Europe. According to David Lord, the head of communications at the European Organization for Nuclear Research (CERN), near

Geneva, the EARN network could be operating within a couple of months. This rapid progress is possible because the technology is identical to the existing BITNET system, and is essentially simple.

There are two essential reasons for connecting computers, says Lord: to control one computer with another (to make one act as a terminal to the other); and to transfer messages and files. Existing telephone-line communications between computers using a simple international communications protocol dubbed X25 coped well with the first use (terminal access). But X25 is clumsy at file transfer.

BITNET and EARN are designed for this latter use — file transfer — which will, for example, enable a researcher in one country to send a research paper or a preprint to a researcher in another, with the minimum of fuss, provided the paper was first produced on-line.

Already the main nodes (junctions) of the EARN network are defined. Rome will be linked to Haifa, to CUNY in New York and to CERN in Geneva. CERN will link in turn to Paris, GSI Darmstadt in West Germany and the Rutherford Appleton Laboratory in Britain which will provide a gateway into the existing British JNT (or Janet) nationwide academic network. Paris will link to Madrid and to Stockholm, a gateway into a Nordic network.

At present, Britain seems to be best provided with academic networking, with Germany rising fast in second place. Germany is planning to provide 10 internal (national) access points to EARN.

But why has Europe left it to a US computer company to establish such a network? This is a difficult question to answer. The information directorate of the European Commission in Luxembourg does in fact run a substantial network, Euronet Dyane, but in practice this has been developed mostly to provide terminal access to various databases and database hosts in Europe. And since databases are mostly abstracting services of most use to librarians, Euronet Dyane has developed more towards library services than to linkage between researchers.

Meanwhile, though, the Commission is developing programmes such as the European Programme of Research in Information Technology (Esprit). In such programmes, new standards for inter-computer communication are being developed — standards that will be more powerful than BITNET and EARN. This is the right job for the Commission, Lord thinks. But meanwhile users in urgent need of communications should not ignore *ad hoc* offers like IBM's EARN. The danger to Europe and the benefit to IBM is that IBM standards will be deep-rooted by the time Europe decides on its own protocols.

Robert Walgate

Toxin weapons

Softening of US charges

Washington

THE US State Department may be backing off from its charges of Soviet use of chemical and toxin weapons in Afghanistan and South-East Asia. In a report submitted to the United Nations last week, the State Department said that in 1983 — for the first time since it began levelling charges of Soviet use of chemical weapons three years ago — it could find no confirmed evidence of chemical attacks in Afghanistan. And in Laos and Kampuchea, where the State Department has accused Soviet-backed government forces of attacking resistance fighters with "yellow rain", supposedly a mixture of trichothecene mycotoxins derived from fungi, the report said that the "lethality of attacks" had decreased in 1983.

Only one of more than 200 environmental and biological samples collected in all three areas in 1983 proved positive for the presence of these toxins. The State Department apparently based its allegations of continued attacks in Laos and Kampuchea on intelligence reports, eyewitness accounts and medical data on symptoms suffered by victims of the alleged attacks.

While maintaining that the United States had "strong evidence" to support its earlier charges, the report seemed at least to open the door to some explanations that stop short of toxin-weapon use, which is a violation of international treaties. "Some deaths associated with toxic attacks occurring in 1983 resulted from secondary effects", the report acknowledged, "such as eating contaminated animal products after an attack."

Critics of the State Department's earlier reports have suggested that use of non-lethal agents such as tear gas may explain the symptoms reported and the presence of gas masks and other material among Soviet forces in Afghanistan. Tear gas is not considered by the United States to be a chemical weapon under the 1925 Geneva Protocol banning chemical warfare.

Other critics, notably Matthew Meselson of Harvard University, have reported strong evidence for a natural explanation of the illnesses reported. Meselson said last week that he is now firmly convinced that the "yellow rain" samples collected by the State Department are in fact pollen-laden bee droppings, perhaps naturally contaminated by toxin-producing fungi.

Meselson, who called last week's report a "retreat" from the State Department's earlier accusations, was particularly critical of the government's continuing failure to publish its findings in an appropriate scientific format, with full details of when and where the samples were collected and of how control samples were chosen.

Stephen Budiansky

Up and away

Tokyo

ALTITUDE control and telemetry tests of the Exos-C satellite, launched last week by Japan's Institute of Space and Astronautical Science (ISAS) from the Kagoshima Space Centre, have proved successful and the first scientific data are expected to be transmitted back in the next few days.

Exos-C, given the Japanese name Oozora (sky), is a part of Japan's contribution to the Middle Atmosphere Program (MAP). It is the ninth scientific satellite to be successfully launched by ISAS, the academic space institute run by the universities, in a space programme totally independent of Japan's commercial space organizations. Launched into a low elliptical orbit (perigee 354 k, apogee 865 k, inclination 74.6 degree, period 96.9 minutes) by an MU 3S rocket, the satellite carries a range of equipment designed to measure minor constituents of the middle atmosphere — carbon dioxide, ozone and aerosols, including those of volcanic origin — from all regions of the globe. Data on ionospheric plasma conditions will also be collected.

The principal instruments on board are a scanning 1.2- μ m infrared radiometer for measuring ozone density at 70–90 km altitude, a back-scatter ultraviolet spectrometer which will give the ozone profile at 25–60 km, and a solar image radiometer for observing stratospheric aerosols and ozone levels.

Alun Anderson

Israeli scientists

Travel to India prevented

Rehovot

A DOZEN Israeli scientists who had planned to attend last week's seventh International Biotechnology Symposium in New Delhi did not do so because of difficulties placed in the way of their obtaining visas.

Most of them gave up months ago, when the delaying tactics of the organizing committee, headed by Professor T. K. Ghose of the Indian Institute of Technology, discouraged them from carrying on their efforts to attend the meeting. But two scientists who were to make presentations in Delhi — Professor Jonathan Gressel of the Weizmann Institute and Dr Stefen Rokem of the Hebrew University — kept trying to get there until the departure of the last flight before the symposium.

Gressel first wrote to the Indian organizers on 21 March 1983, requesting that they send him application forms for a visa because there is no Indian mission in Israel. Five months later, Dr S. Chand advised him that the organizers had arranged "for the concerned authorities to provide a landing permit" on his arrival in India. But

Alitalia, which operates the only scheduled air service between Israel and India, informed Gressel that he could not board one of their planes bound for India on the basis of such a letter without details about the visa.

This should have been simple to arrange because according to Alitalia, the Indian authorities routinely provide such details where ordinary Israeli tourists are concerned. In any case, Gressel was finally informed on 21 January 1984 that he would have to go to the Indian Embassy in Rome for a visa. This he refused to do because it was contrary to the original promise, would cost him US\$200 more for an additional stopover and, as he cabled Chand, "the Indian embassies concerned usually claimed to have no record of their request".

This is only one of many instances where Israeli researchers have been prevented from attending international meetings in India, but no less troubling, says Dr Baruch Eyal, head of the Division of International Scientific Relations in the Israel National Council for Research and Development, is the fact that many Indian scientists invited to take part in meetings in Israel were prevented by their government from doing so.

Meanwhile, Professor Gressel and Dr Rokem are waiting for some action — even if only after the fact — from Professor H. G. Schlegel, chairman of the International Committee on Economic and Applied Microbiology, a sponsor of the symposium.

Nechemia Meyers

UK science

Declining and parasitic

SIR Hermann Bondi, chairman of the Natural Environment Research Council, warned this week that British science is in decline and that without substantial transfers of resources British environmental scientists will be unable to play an international role. Sir Hermann was thankful that the government had maintained a level science budget within its overall public expenditure policy. But he pointed out that many of Britain's competitors are increasing their capabilities while British science "is standing still on an escalator that is going down".

Sir Hermann argued that cooperation between ships and satellites is essential for the future development of marine research. But, he said, at present there is no money for oceanographers, geologists or ecologists to contribute instrumentation for satellites. All they can do is to analyse information from others' instruments; and "I do not believe that a situation where the scientist depends on the kindness of others is one that is tenable for much longer".

The National Environment Research Council was recently told by a House of Lords Select Committee formally to liaise with the Science and Engineering Research Council on basic research. Some joint projects concerned with remote sensing (see *Nature* 16 February, p.586) have been in difficulties because of questions about areas of responsibility — and, of course, financial support.

Tim Beardsley



A NEWLY discovered Roman fort in Scotland, one of three revealed by aerial photography during the exceptionally dry summer of 1983. The fort is on the South Esk immediately to the north-east of Kirriemuir, near the castle of Inverquhar, and was built there in the Flavian period, probably by Julius Agricola who, as governor of Britain from AD 83–87, instigated a campaign to contain the Caledonian tribes. For the Royal Commission on the Ancient and Historical Monuments of Scotland, 1983 was a vintage year, with the discovery of new examples of 5,000 year-old long timber halls, whole cemeteries of burial cairns and Iron Age villages, in addition to the forts. □

Nuclear waste rules undecided

DELEGATES at the London Dumping Convention last week failed to reach a decision on whether the "emplacement" of high-level radioactive wastes under the sea bed comes within its purview.

The issue has become controversial because it is claimed that recent research shows that loading wastes into torpedoes that would penetrate deep beneath the sea bed is at least feasible. But disposal at sea of high-level waste is specifically outlawed by the convention: hence the legal argument which occupied a good part of last week over whether the sea bed includes the earth below the sea bed. Nations which are producing high-level wastes figured prominently among those who held that sub-sea bed emplacement is outside the terms of the convention, while Spain and some Scandinavian countries were notably inclined to the opposite view. The time-honoured solution eventually reached was that a decision should be put off until the convention meets again next year, when pollution by plastics will also be considered.

A year ago, the convention adopted a

resolution calling for a temporary halt to dumping of all radioactive wastes until such time as the consequences have been fully explored by a scientific working group, and since then the legal niceties of the question have been scrutinized. Last week it was eventually agreed that the group would be drawn from a panel of impartial scientific experts drawn up by the International Council of Scientific Unions. The group's recommendations will be made and examined by representatives of the 53 contracting parties to the convention before next year's meeting. As this is likely, in the words of the convention, to be an "onerous task", next year's meeting has been postponed by six months until August 1985. This means, incidentally, that the moratorium on even low-level dumping is also effectively extended, a fact which caused some joy to environmental groups. Britain is opposed to, and has never recognized, the moratorium but the matter has become somewhat academic since industrial action by the National Union of Seamen has prevented Britain's annual dump taking place.

Tim Beardsley

Biotechnology

Interferon patent contested

Boston and London

BIOMED NV has scored a provisional victory in the competition to market alpha (leukocyte) interferon. The European Patent Office has formally notified the biotechnology company, based in Cambridge, Massachusetts, and Geneva, Switzerland, that it will receive a product patent for its alpha interferon synthesized by recombinant DNA technology. The patent could bring Biogen and its licensee, Schering-Plough Corporation, commercial dominance in selling alpha interferon (trademarked Intron) in the important European pharmaceutical market.

Genentech, the South San Francisco biotechnology company, is crying foul, and plans vigorously to challenge the European Patent Office's action during the nine month appeals period after the patent's issue, which may yet be some months away. According to Genentech and its partner F. Hoffman LaRoche and Co., it would be wrong for Biogen to receive a patent covering production of any cloned alpha-interferon in bacteria, yeast or mammalian cells.

Biogen's patent application refers to recombinant DNA molecules and processes for the production of human interferon alpha-like polypeptides. The initial research, carried out at the University of Zurich, led to a paper in *Nature* (284, 316; 1980) describing the production in *Escherichia coli* of a polypeptide with interferon activity. But according to Genentech the patent related to this work should only cover the cloning and production of a precursor to alpha interferon, not

the mature interferon that is biologically cleaved from it. Genentech's work was published in *Nature* six months after Biogen's (287, 411; 1980) and its patent application, filed six months after Biogen's, claims priority for producing mature interferon.

Biogen notes that this will be the first patent issued for a commercially significant product from the biotechnology industry. It is aggressively seeking issue of a US patent, and plans quickly to enforce its exclusive rights to European manufacture and sales. Biogen vice-president Peter Feinstein says that the company will soon inform its competitors that they must cease European production, although it is unlikely that any notice will be taken of such a demand.

Schering-Plough has sought regulatory approval for two uses of alpha interferon in cancer therapy. Last year multiple myeloma, malignant melanoma and Kaposi's sarcoma were among the conditions where it was said to have therapeutic potential. The extensive clinical trials and regulatory hurdles that must be passed for approval of new cancer treatments will probably prevent the drug — no matter who is making it — from reaching the market for at least another year or two.

The immediate beneficiaries of alpha interferon will be patent lawyers, who are now rubbing their hands with glee at the prospect of what is likely to be a prolonged case. And with Biogen also seeking patent protection for beta and gamma interferons, there is likely to be much more to come. **Christopher Earl & Tim Beardsley**

More Antarctic chill for Japan

Tokyo

JAPAN is to increase its stake in Antarctica with the building of a third permanent research station for the national polar research institute (NIPR). The go-ahead for the new base came directly from Antarctica after ten members of the 25th summer Antarctic expedition, accompanied by NIPR head Takeshi Nagata, reached the proposed site in four snowmobiles and a helicopter last week.

The new research station will be built close to the now abandoned Belgian base of Roi Baudouin at the foot of 3,180 Mt Ser Rondane (71.3S, 24.4E), 100 km inland from Breid Bay and 650 km from Japan's large Showa base on Ongul Island, Lutzow Holm Bay. Weather permitting, use of the base will begin in the next Antarctic summer. Seven or eight scientists will soon be able to overwinter, adding to the fifty-five Japanese already in permanent occupation at the Showa and smaller Mizuno bases.

Funds for the new station will be provided by an increase of NIPR's annual

budget to 3,300 million yen (\$14 million). Tatsuro Matsuda, head of NIPR's research division, stressed that the new station, like those already established, will be used purely for academic research. Meteorology, glaciology, geology and cartography will be the principal areas of investigation, together with participation in the auroral and middle atmosphere study programme for which rockets and balloons are launched from the Showa base. Despite increasing pressure from the commercial exploitation of Antarctica, NIPR has no intention of undertaking any estimates of potential mineral reserves, said Dr Matsuda.

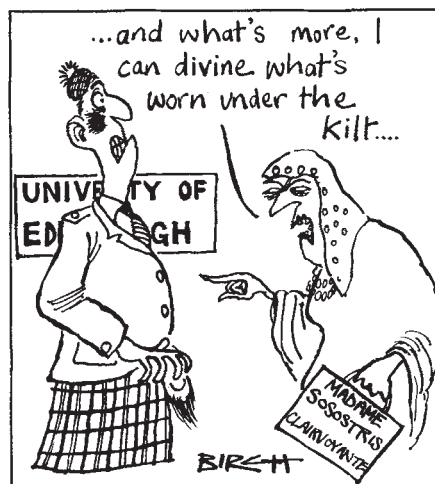
Within Japan itself, most pressure for the commercial exploitation of Antarctic resources comes from the fishing industry, now by far the largest in the world. As two-thirds of the catch comes from foreign fishing grounds, many of them within other nations' 200-mile economic zones, the advantages of exploiting Antarctic krill are great. **Alun Anderson**

Paranormality

Koestler chair for Edinburgh

CLAIRVOYANTS, spoonbenders and others boasting paranormal talents will shortly be needed for research to be conducted at the University of Edinburgh. The university is to become the first in Britain with a professorial chair in parapsychology, and the successful applicant could be appointed as early as September.

The chair is supported from the estate of the late Arthur Koestler, who committed suicide with his wife last year (see *Nature* 302, 93; 1983). Koestler became increasingly preoccupied with mystical matters toward the end of his life, and left a sum



believed to be in excess of £500,000 to make permanent provision for a chair. The Koestler Foundation has pledged the interest from a similar sum to support a research programme.

Executors of Koestler's will had to choose between three serious contenders for the endowment — Edinburgh, City University, London, and the University of Wales. Edinburgh was chosen, according to Dr John Beloff, one of Koestler's executors and a member of the psychology department at Edinburgh, because it was able to offer the best facilities: several rooms will be made available for the new professor in the psychology department. But Beloff will not himself be applying for the post, although he does carry out research in the field.

Announcing the establishment of the chair, Dr John Burnett, principal of the University of Edinburgh, said that after widespread consultation it was clear there was overwhelming support within the university for the idea. The new professor will carry out research into "the capacity attributed to some individuals to interact with their external environment by means other than the recognized channels". Dr Burnett said that the investigations were likely to lead to new knowledge, whatever their outcome.

Tim Beardsley

Diablo Canyon safe

SIR — In the news article "Disaster by many small cuts" (*Nature* 306, 631; 1983), the following statement was made about the Pacific Gas and Electric Company's nuclear power plant at Diablo Canyon. "In California the two reactors at Diablo Canyon have only just received permission to load fuel after a year's delay caused by the last-minute discovery that the reactors, built above the San Andreas fault, would not be able to meet NRC requirements for withstanding earth tremors."

In fact, however, the Diablo Canyon power plant is located 60 km from the San Andreas fault and its safety in the event of a magnitude 8+ earthquake on this fault has never been questioned.

On the nuclear power stage, there are three principal actors: the power company that wishes to construct the plant, the Nuclear Regulatory Commission that acts as referee and the organized opposition that does not want a power plant built at that location. The statements of these actors are widely reported by newspapers, especially those issued by the opposition. Unfortunately, the views of the engineers who design nuclear power plants and the views of the informed engineering community are neither reported nor sought by the news media. Consequently, much misinformation has been implanted in the public consciousness. G.W. HOUSNER

*Division of Engineering
and Applied Science,
California Institute of Technology,
Pasadena, California 91125, USA*

Nuclear Phoenix

SIR — Contrary to the report of Stephen Budiansky (*Nature* 19 January p.201), the AIM-54 Phoenix is not new but has been manufactured since 1973. It is deployed by the US Navy on the F-14 Tomcat (the US Navy's carrier-borne long-range interceptor), and has been since the aircraft was first delivered in September 1973.

The Phoenix is very well known because of the large number of records it holds, so the fact that the report is approximately ten years out of date is worrying. Is this due to a singular although very serious mistake or does it indicate sloppiness in the preparation of the original manuscript? I hope it is the former because the general public is appallingly informed about the nature and roles of nuclear weapons and there is a real need for properly informed debate in the general population rather than the rabble-rousing that seems to be the present case. J.A.A. CAMBERS

*Max-Planck-Institut für
Molekulare Genetik,
Innistrasse 63-73,
1000 Berlin 33 (Dahlem), FRG*

● Stephen Budiansky writes: What is new is the development of a nuclear warhead for this missile. □

In the abstract

SIR — A correspondent recently complained about the citing of abstracts of conference presentations as literature references, primarily on the basis that they are not properly refereed. True, but the problem does not stop there. For instance, in many cases, the abstracts bear little resemblance to the paper eventually presented at the conference. In clinical papers in particular, for instance, the number of patients involved in a study as reported in the abstract may be entirely different from the number reported at the conference.

A greater problem is that many abstracts that appear in the proceedings of a meeting may actually represent papers that were never presented. This is especially true of poster presentations. At a recent conference on food technology, I found that only about 40 per cent of the posters listed in the abstracts were actually presented. Papers from Eastern Europe frequently encounter this fate because scientists for whatever reason may find it difficult to follow up their preparation of an abstract by attending the conference.

Worst of all, a paper that appears on the programme and in abstract form is sometimes withdrawn at the last moment. This happened recently at a leukaemia conference where, on the basis of the abstract, I waited for two days to hear what seemed to be one of the more interesting papers at the conference. The paper was never presented.

These criticisms apply less to small abstracts with little hard information than they do to more extensive abstracts, sometimes of full paper length. Although the latter abstracts may be very detailed, they mean very little if the paper is not presented, or if the presenter strays widely from the original script.

Abstracts may be a good guide to what is going on at a conference, but it is laughable to consider them as anything more than a remote approximation of scientific reality.

JOHN F. HENAHAN

*118 Ballinclea Heights,
Killiney, Co. Dublin, Ireland*

Biophysics

SIR — The leading article on biology as part of physics encourages me to make a general remark on holistic versus reductionist approaches in physics. Since the number of microstates increases exponentially with the number of particles, N , say 10^N , even a super computer could, in the life time of the Universe, solve only for $N \approx 20$ unless systematic ways of finding them exist. (Compare my article on the connection of macro and micro physics, *Riv. nuovo Cimento* 3, 490; 1973.) Macroscopic equations, such as the Navier-Stokes equations, can, however, be derived from macro-physics without actually solving the N -body problem. This requires the definition of appropriate macro-concepts, such

as current density, in terms of micro-concepts.

When one is far from equilibrium, as in biology, a basic problem arises — namely to find the relevant macro-concepts. Structure alone is insufficient as active systems are highly excited, and hence new macro properties arise which cannot be found by systematic investigation of the N -body problem. From this viewpoint, the holistic and reductionist approaches supplement rather than contradict each other.

H. FROHLICH

*Department of Physics,
University of Liverpool,
Oliver Lodge Laboratory,
Liverpool L69 3BX, UK*

Creative paradigm

SIR — The article on creationism in Texas (*Nature* 306, 528; 1983) points to what may be a more general problem in basic education. Compared with the difficult choices on complex issues that students will have to make when they are adults, the question of human origins is benign. So instead of simply absorbing "facts" in the classroom, should not students be learning how to evaluate information derived from different sources, taking into account the reliability of the source, the accuracy of the information and the consequences of their misinterpretation of the material?

This skill must be taught and encouraged; it requires a good working knowledge of language, particularly the ability to read a passage accurately and critically. Let creationism and evolution be discussed if that is what teachers, students and parents desire, but present the issues in such a way that students learn the difficulties inherent in evaluating historical material. For example, originally the Bible was not written in English or any other contemporary language in common use, and so we must question whether the authors intended to say exactly what has reached us after centuries of translation and revision.

Why not use this controversial issue as a way of training students in the proper evaluation of facts and ideas? Any idea worth being taught and passed on to the next generation ought to be able to stand up to comparison and criticism, and unless students learn this in school, they will become adults who are ready to believe anything. J. POTTER

*Strawberry Hill,
Bar Harbor, Maine 04609, USA*

Lesson in objectivity

SIR — "Just how objective is science" by N.S. Hetherington (*Nature* 306; 727 1983) was interesting and objective. However, I think science is objective by definition. I would have entitled the article: "Just how objective are scientists?"

J. M. GOLDSCHVARTZ

*Clavecimbellaan 273,
2287 VK Rijswijk zh, The Netherlands*

What economists have to say

Jeremy Bray*

The newly-appointed Labour opposition spokesman on science and technology in the British House of Commons argues that scientific method itself is the most important contribution that the science of economics has to make to government.

MY message, to echo Kennedy's Inaugural, is, "Ask not what government can do for science. Ask what science can do for government." While science and technology can usefully give the cornucopia of material goods a further shake, their more difficult and more rewarding task is the contributing of ideas.

The most important contribution that science has to make to government is scientific method itself. Take food policy as an example. The indications are that a reduction of salt, sugar and saturated fatty acids and an increase of dietary fibre in the diet of the nation would substantially reduce mortality and improve health in the middle years of life. The evidence is sufficient to have persuaded the better informed to take account of it in their own diet. Mortality and ill health is worse in working class areas, which are more exposed to the pressures of the market place.

The government could wait (as it has tended to do over smoking and lung cancer) until the last "i" is dotted in the scientific evidence. Or it could say, "This is the balance of the evidence. We recommend — but will not of course enforce — these dietary goals. We are taking these measures on food labelling, availability of processed foods, meat grading, animal and plant breeding research, public catering, food and agricultural subsidies, and health education. We shall monitor diet and health, and step up medical research in the field. We shall consult and inform the public fully as we go."

The former course is to treat scientists as oracles. The latter is to involve the public in scientific method. That it is also more democratic and responsible government is no accident.

As another example, take nuclear disarmament. Consider the following result in the theory of dynamic games, produced recently by the mathematician Steve Smale. Two prisoners in a cell can each play easy or tough. If one plays tough, and the other plays easy, the first scores 3 and the second 0. If they both play easy they each score 2. If they both play tough they score 1. There is no interesting solution of the single play prisoner's dilemma. But in the repeated play Smale defines a "good strategy" as one in which one player creates a situation in which the simple self-interest of the other is also to follow a "good

strategy", and the average score converges on 2-all. One such good strategy is to play easy if your average score is over 2, and to play tough if it is under 1. Between 1 and 2 you play easy if your average score is higher than the other player's, or only a little bit lower. (Steve Smale, "The Prisoner's Dilemma and Dynamical Systems associated to Non-Cooperative Games", *Econometrica* 48, No. 7, 1617-34; 1980). No such logic has been presented in Britain for the deployment of Trident, Cruise or Pershing missiles. And it demonstrates neatly the inevitability of unilateral actions in a multilateral situation. The choice is between unilateral actions which escalate or de-escalate the confrontation.

The field of government with the widest need for scientific method is economic policy. It is not necessary to argue whether or not economics is a science. It is sufficient to argue that the economics used by government has been falsified science.

Like so much else in science, economics is in a healthy state. The economy and often the teaching of economics are not. But in micro-economic theory, the theory of the interaction of individual economic agents is well set out mathematically. The result is not an atomic theory of the behaviour of individuals on the model of the old physics, on which a kinetic theory and properties of matter can be built, but rather a means of describing and exploring the very different kinds of behaviour in the large that emerge from different behaviour of individuals.

Much current interest focuses on behaviour in which the individual is influenced not merely by present conditions, but by his expectations of future conditions, formed from a true model of the system. There is the highbrow theory which looks at how individuals learn to form their expectations in this way. And there is the middlebrow theory which looks at the behaviour of the system when they have learned.

But now suppose the individual is not just maximizing his own objective function with an adaptive stochastic control algorithm. If he is watching the strategy of other individuals, and adapting his algorithm in the manner of the Smale result, and other individuals are doing the same, what kind of meta-game results? Furthermore, since individuals in fact act by rather simple rules of thumb, what rules of thumb will approximate to this

behaviour? And what are the rules of thumb and the properties of the system, which produce the tendency which we observe so often for the system spontaneously to develop progressively higher levels of organization, in apparent contradiction of the second law of thermodynamics? What makes behaviour flip spontaneously from one pattern of organization to another? Then, to make a huge leap, what are the influences that can be brought to bear in national policies which will make the flips benign, within and between nations?

In fact it is not a leap. There is much work going on in between of equal interest and merit in its own right. The macro-economists, working generally with 2 to 10 equation theoretical models of national economies, have paralleled the work of the microeconomists, exploring the properties and policies of closed and open economies, with capital flows as well as trade, forward looking expectations, limited information and uncertainty.

At the empirical stage, econometrics has become a sophisticated branch of applied statistics over the past 20 years because of the difficulty of drawing statistically sound conclusions from the short samples of data with high measurement errors generated by the uncertain and changing relationships inherent in economics.

The balance of evidence of all this work, micro and macro, theoretical and empirical, has been effectively ignored by the government.

Following our work and representations on the Treasury and Civil Service Select Committee in the last parliament, the Social Science Research Council (now the Economic and Social Research Council), together with the Treasury and the Bank of England, set up a Macroeconomic Centre at the University of Warwick to test and compare the large empirical models and the forecasts produced from them.

The latest study of Friedman's work carried out recently by Hendry and Ericsson for the Bank of England Panel of Academic Consultants (Panel Paper No. 22) is not an anti-monetarist polemic. They expose Friedman's errors in seeking to analyse complex stochastic processes while eschewing modern econometric methods. He has not established any sense in which the velocity of circulation of money is constant, nor has he established that setting a target trajectory for the quantity of money

*Jeremy Bray first entered the House of Commons in 1962 and served in the Labour Government of 1964-70.

is a satisfactory basis for policy.

When the Chancellor of the Exchequer, Mr Nigel Lawson, was asked in the Treasury and Civil Service Select Committee (Minutes, 19 December 1983) for his response to Hendry's criticism of Friedman, he replied that the questioner's belief that the government's policies are based on the writings of Friedman was "curiously unhistoric". They represented, he said, "the prevailing approach during the period of most of the growth that has occurred in the Western world". A minister of health faced with an epidemic would not boast that he was tackling it by the public health methods which had stood Britain in such good stead for the past 200 years, including the period of fastest population growth.

When the intellectual history of economic policy-making in the period 1976-83 comes to be written, it will be seen as a scientific scandal comparable with the Lysenko controversy, but more damaging in its effects. In the Soviet Union, Lysenko did useful work shortening the ripening period of cultivated plants by temperature influences at the start of germination. But he built on it a theory that it was possible to change the heredity of plants, sharply attacking the biology of the West. It was supported by Zhdanov, the best educated member of the politburo, and endorsed pusillanimously by the Soviet Academy of Sciences. Dissenting scientists were removed from research and in some cases arrested. We do these things differently in Britain.

Technical economics has become highly mathematical, requiring a good grasp of linear algebra, analysis, set theory, calculus, convexity, maximization subject to constraints, dynamical systems, probability, statistics, stochastic processes, econometrics and control theory. And those are just the basic tools. *The Times*, the *Financial Times* and *The Economist* do not employ economic correspondents who can read technical economics, so it is never explained to the lay public. It is as if no science correspondent had ever solved Maxwell's equations or sequenced a DNA molecule.

Nor is the economics profession itself always able to cope with the mathematics. It is not just the older generation. There is a vested interest in keeping the mathematical standards for entry to economics departments low to keep up student numbers. Most schools, unaware that economics has become so mathematical, encourage mathematically gifted students to go into the physical sciences. Most universities here and abroad do not have enough staff able to teach good undergraduate courses in technical economics. When vacancies do arise there is sometimes a majority on the appointments committee organized to keep out technically competent economists to avoid showing up less competent present members of the faculty. The cry to appoint a Marxist development economist rather than a mathematically trained economist has gone up recently in more than one

university. It would have attracted the withering scorn of Marx himself, and it is a betrayal of developing countries that they should be sold inferior techniques.

Research budgets have been cut instead of being increased to tackle the problems that needed to be resolved. Governments dodged serious well formulated technical questions about their own policy, and the nation reaped the consequences.

Against this background, in the seventies, the old Keynesian establishment had no politically workable answer to inflation. Its members knew nothing of the techniques and results of modern technical economics. When the challenge of Friedman and his generation of monetarists came along, with its strong political implications, they attacked it with the bows and arrows of their generation. The younger generation of monetarists were no less politically motivated than their elders, but they put forward technically interesting and important ideas such as rational expectations and private agents' responses to changes in government policy. This was getting out of the depth of the more literary-minded monetarists. It was absorbed more quickly by the wider more eclectic school of technically trained economists such as Miller and Buiter, returning to Britain from the US graduate schools at Yale, Princeton, Harvard, MIT and Stanford. Now we have a generation of economists well trained technically, who between them cover the whole range from theoretical micro to applied macro and industrial.

There have been parallel developments in the United States and in the very important area of international economics. In the research and international departments of the Federal Reserve Board in Washington, at Princeton University, the University of Pennsylvania and elsewhere in the United States, and in Japan, preparatory work on international economic cooperation and the international monetary system is going ahead in anticipation of the return of a political climate favourable to cooperation in the United States, Europe and Japan.

After the scandal of monetarism, and in Britain the debacle of the old Keynesians, we are ready to pick ourselves up and dust each other down. The scientific community which is best equipped to understand the technical work of its economist members can help in the exposure of past errors and the design of future policy.

Lord Flowers has expressed the hope that abundant computer power will make technology more accessible to the mathematically untrained (Keynote Address to Second World Conference on Continuing Engineering Education, Paris, April 1983). Whether or not that is likely in the physical sciences and engineering, it is happening in technical economics. By 1988 and the next election, inexpensive publicly available microcomputer models of business enterprises and public authorities will be linked to national economy models,

if not by the enterprises themselves, then by their employees, trade unions and investment analysts. The macro-modellers wishing to survive will be incorporating the feedback. They will then be presenting, not mirages after the manner of current forecasts of the unforecastable, but a channel of communication, cooperation and competition between independent but interacting agents. The possible performance of such a system is of course superior to that of purely competitive games conducted in ignorance of the technology and plans of other players.

The shape of the task which we face in returning to full employment is best set out in the industry and occupation projections of the Institute for Employment Research under Professor Robert Lindley at the University of Warwick (Review of the Economy and Employment, Summer 1983). Working on the assumption of the continuation of present government policies, unemployment is not expected to fall much below 3 million by 1990, with continuing job losses in primary and manufacturing industries, balanced by increases in professional and miscellaneous services, construction, and social and public administration. The decline in oil production in the North Sea will cause balance-of-payments difficulties by 1990, restricting the competitiveness of British industry. We shall probably have to expand manufacturing industry to contribute to the balance of payments. But it is unlikely to create many jobs.

To restore growth and full employment we shall need a much greater effort on the balance of payments in manufactured goods, since it will deteriorate in oil and have limited possibilities for improvement in invisibles. The greater the improvement in manufactures, the more will the balance-of-payments constraint allow us to increase employment in services and construction.

Most of the loss in manufacturing output since 1979 cannot be restored now simply by increasing demand because the capacity has been scrapped. Import controls offer only temporary relief if they are not to produce a permanent loss in real incomes. So the increase in manufacturing exports and import substitutes will have to be in world competitive products from new capacity.

As Pavitt has argued (*Papers in Science, Technology and Public Policy*, No. 3, 1983), the realization of competitive advantages at the national level depends on building company by company on their successful directions of development within a generally high level of research and development and investment activity. It is not a matter of picking winners out of the blue, or merely in terms of broad sectors. Nor is it a matter of broadcasting largesse indiscriminately.

The present government will not use the powerful information system now developing for economic management purposes on grounds of principle. I hope the scientific community will. □

Nuclear winter not yet established

Talk of some of the consequences of nuclear warfare had better be postponed until the underlying assumptions are better understood.

WHEN all the bombs go off, will we frizzle or be frozen? Or, more accurately, can those who survive the first fate hope also to escape the second? These are among the questions raised by the renewed interest among geophysicists in the climatic consequences of several nuclear explosions. It is too soon to answer them with any clarity.

That the surface temperature of the Earth would be reduced if the atmosphere of the Earth were laden with dust and aerosol after a nuclear war is not, as a general proposition, open to dispute. Substantial amounts of dust in the atmosphere would alter the natural balance of radiation near the Earth's surface in two ways, by preventing solar radiation reaching the surface and by depositing energy in those regions of the atmosphere where the dust lies, changing the hydrodynamic behaviour of the atmosphere and thus the weather in the process. Developments such as these are responsible for the climatic consequences of volcanic eruptions and have also been invoked (in conjunction with the arrival of meteoric objects) to account for species extinction at the end of the Cretaceous (see *Science* 208, 1195; 1980).

Difficulties arise only when attempts are made to calculate the climatic consequences of atmospheric dust. Even the empirical correlation of the magnitude of volcanic eruption and their climatic after-effects is made fuzzy by the variability of the character of the dust and aerosol injected into the atmosphere (see *Nature* 307, 107; 1984). By expectation and empirically, the size distribution of the ejected material is important for two reasons — smaller particles will remain in the atmosphere for longer, while their larger specific surface area will more directly influence the radiation balance.

Much of the recent interest in the recalculation of the climatic consequences of nuclear war has been stimulated by measurements of dust particles from eruptions of volcanoes such as Mount St Helens and El Chichon, which have suggested that fine material may be unexpectedly conspicuous. The topic was aired at a public meeting in Washington last autumn and also at a symposium at the meeting of the American Geophysical Union at Los Angeles in December, but the only published account of the calculations so far is that of Professor Carl Sagan and his associates (see *Science* 222, 1283; 1983). A committee of the US National Academy

of Sciences is due to report on the subject later in the year. Meanwhile, the article by Covey, Schneider and Thompson (this issue, p.123) should be regarded as but the latest contribution to what seems certain to be a prolonged and contentious argument.

The Sagan calculation, acknowledged by its authors to be preliminary, turns on the assumptions made about the input of dust and aerosol into the atmosphere in the aftermath of a nuclear war, but otherwise involves a simple model of the energy balance within the atmosphere. That document is less than convincing for two reasons — the promised detailed discussion of the assumptions remains unpublished, while the pardonable simplicity of the calculation of climatic effects, innocent as it is of the feedback mechanisms likely to occur in the real atmosphere, is likely to exaggerate the severity of what is called the nuclear winter. What Covey *et al.* have now done usefully complements this earlier calculation, but only on the second of its weak points.

The new model of the Earth's atmosphere is nevertheless sophisticated enough to deal with at least low-frequency variations of climatic quantities with both latitude and longitude, while the atmosphere is represented vertically by no fewer than nine successive slices. Moreover, the new climatic model breaks new ground by means of its allowance for the occurrence of cloudiness in the lower troposphere. The failure to include such effects has been one of the weaknesses of attempts to calculate the climatic consequences of, say, the accumulation of carbon dioxide in the Earth's atmosphere. It is not clear whether the new model will successfully accommodate these complications, and it may even be thought unfortunate that it has been applied to the calculation of the nuclear winter before much has been published of its usefulness in more conventional calculations.

With these reservations, the new calculations do soften the results described by Sagan *et al.* in the expected direction. Covey *et al.* explain that they have arranged in their model calculation that, in middle latitudes, solar radiation should be predominantly absorbed by atmospheric dust. Virtually none of it will reach to the surface in middle latitudes. The conclusion is that the grip of the nuclear winter will be strongest on the land masses of the Northern Hemisphere (where, it is supposed, the bombs will go off). Some may think it

ironic that the territory of the superpowers will be most affected, although this is a consequence of the assumptions made about the relative specific heat of land and sea.

The assumptions about the input of atmospheric dust are, by contrast, no more persuasive than those of Sagan *et al.* Faced with the understandable difficulty of predicting the course of a nuclear war, Covey *et al.* conservatively neglect the effects of dust kicked up mechanically, much of which is known from test explosions to be carried into the stratosphere. Instead, they confine their attention to the smoke from postwar fires, which they assume will be enough virtually to make the atmosphere opaque to solar radiation. For want of a technique for predicting what might actually occur, they suppose that the smoke will remain in place, neither spreading nor settling, during the twenty days spanned by their calculations. Correctly, the authors draw attention to this limitation of their work. They do not say explicitly, no doubt because the point is obvious, that their calculated winter must be an exaggeration.

The result is that while the new calculation shows that the grip of the nuclear winter will vary from place to place much as does the weather, there is still a long way to go before its intensity can be calculated. That unremarkable conclusion, at this early stage in the development of climatic models and in the face of prevailing ignorance about the likely behaviour of large amounts of atmosphere dust, should cause no surprise. If, on the basis of calculations so far published, some people should refuse to believe that there would be a long winter after a nuclear war, they cannot easily be refuted.

Reasonable people will no doubt prefer to follow an intermediate course. Because some parts of the world's population would not be directly affected by a nuclear war, it is not simply an academic matter to consider what the climatic consequences would be. Moreover, as Covey *et al.* point out, the problem is interesting. In time, it should also be tractable. Until then, however, there is the strongest case for asking that the prospect of a nuclear winter should not be made into a more substantial bogeyman than it is by those who earnestly wish to avert the prospect of nuclear war as such. By clouding the case with disputable predictions, they are in danger of weakening it.

John Maddox

Immunology

The immunoglobulin superfamily takes shape

from Alan F. Williams

In this issue of *Nature* on page 37, Mostov, Friedlander and Blobel publish the primary structure of a receptor (the poly Ig receptor) for polymeric immunoglobulin A (IgA) and immunoglobulin M (IgM)¹. Surprisingly, the receptor is structurally related to immunoglobulins, in particular their variable (V) domains. Thus immunoglobulins and their receptor are members of the same superfamily of molecules.

Cell-surface receptors for immunoglobulins play a vital part in immunity by transporting immunoglobulins or by triggering immune effector functions after antibodies have bound to antigens. They include the IgG receptors of macrophages and antibody-dependent cytotoxic cells, the IgE receptor of mast cells, a receptor for transport of IgG across the placenta and the poly Ig receptor for IgA and IgM. In the gut, the poly Ig receptor is first expressed on the surface of the epithelial cell that is exposed to the inside of the villus. There it binds dimeric IgA secreted in this area by plasma cells. The complex of IgA and receptor is internalized into a vesicle, transported across the epithelial cell and, by fusion of the vesicle with the plasma membrane, exposed on the cell surface to the lumen of the gut. The receptor is then cleaved by a protease to release dimeric IgA with about 75 per cent of the receptor attached to it. The released part of the receptor is called 'secretory component'.

Mostov and his colleagues now provide the first full sequence data on any of the immunoglobulin receptors, enabling a model to be drawn for its structure at a cell surface and for that structure to be compared with structures of the immunoglobulin superfamily (Fig. 1). Relationships within this family are based on the presence of domains of sequence containing about 100 amino acids that are folded into two β -sheets stabilized in a sandwich by a conserved disulphide bond² (Fig. 2). The immunoglobulin domains and segments similar to them are shown by the circles in Fig. 1. The poly Ig receptor is postulated to have five segments related to immunoglobulin domains. It is depicted in Fig. 1 looking something like an IgM heavy chain, but in reality is probably folded back on itself with interactions between the immunoglobulin-like domains stabilizing the structure. Another notable feature of the poly Ig receptor structure is the stretch of 103 amino acids postulated to be inside the cell. This cytoplasmic tail, which may be involved in traffic of membrane vesicles, is of impressive size compared with the postulated tails of the other molecules in Fig. 1 (see legend).

It is striking that four of the five domains (labelled V in Fig. 1) of the poly Ig receptor clearly resemble immunoglobulin V domains and Thy-1 antigen³ whereas the fifth domain (labelled C in Fig. 1) is more like an

immunoglobulin constant (C) domain. V domains differ from immunoglobulin C domains by having an extra loop of sequence in the middle of the folded structure (Fig. 2). Domain five is not completely C-like since it has a V-like pattern of sequence around cysteine residue 520 although the length of sequence between the cysteines is more like that of a C domain. Also in this central region, the sequence Pro-Thr-Lys-Leu-Ser-Ser-Gly of the poly Ig receptor can be aligned with Pro-Ala-Val-Leu-Glu-Ser-Ser-Gly of the C_H1 domain of the human IgG1 (ref. 4).

What does all this mean? Could the structural resemblances be due to convergent evolution to a structural type that is favourable for expression at cell surfaces or are they truly homologous? The strongest argument against their convergent evolution is that the enzyme superoxide dismutase has a three-dimensional structure similar to an immunoglobulin domain⁵ yet does not have the conserved disulphide bond characteristic of the immunoglobulin fold or any other detectable sequence resemblance to immunoglobulin⁶. Therefore, most probably, the molecules in Fig. 1 are related in evolution and share the functional property of being involved in recognition events at the cell surface. Certainly immunoglobulins are the B-lymphocyte antigen receptor and major histocompatibility complex antigens in association with the foreign antigen on a presenting cell are what are recognized in T-lymphocyte responses⁷. However, since the poly Ig receptor is not involved in antigen recognition, immunoglobulin-related molecules can obviously have functions other than those immediately involved in the immune response to foreign antigen. The same conclusion can be drawn for Thy-1 antigen since in various species it is expressed on neuronal cells and fibroblasts but not on lymphocytes. The functions of Thy-1 and Qa and Tla antigens (see Fig. 1) are unknown but the identification of mRNA for a probable Qa/Tla antigen in transformed cells and embryonal carcinoma cells has prompted the suggestion that these antigens are targets for natural killer cell cytotoxicity⁸.

An evolutionary origin for the immunoglobulin superfamily has previously been suggested on the basis that recognition in cell interactions might be similar to that occurring between immunoglobulin domains and receptors^{9,10}. We suggested that molecules corresponding to a single immunoglobulin domain were first expressed at cell surfaces at an early stage in the evolution of multicellular organisms and that these molecules interacted with receptors on other cells to mediate cell interactions. Both arms of this recognition system may have duplicated and diverged to provide increasingly sophisticated recognition systems to control the specificity of cell interactions in tissue formation. From these recognition systems the immune recognition molecules could have evolved. Although we discussed this

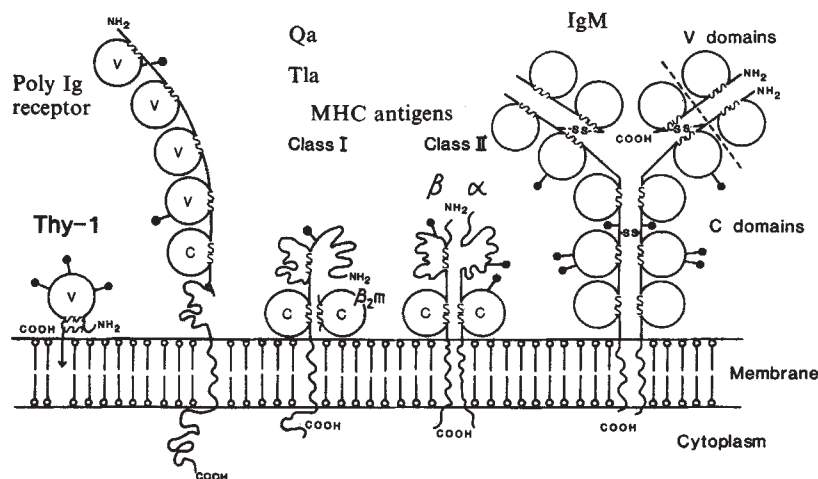


Fig. 1 The immunoglobulin superfamily illustrated at a cell surface. IgM is shown as typical for all immunoglobulins and is the antigen receptor of primary B lymphocytes³. MHC antigens are the major histocompatibility antigens recognized in graft rejection and function as determinants that are recognized along with foreign antigen in T-lymphocyte responses. Data for class I from refs 11, 12; for class II from refs 13, 14. Thy-1, Qa and Tla are all differentiation antigens expressed on various cell types and are of unknown function. Data for Thy-1 from ref. 15; for Qa and Tla from ref. 12. Ig domains and their homologous sections in other molecules are shown by circles and the letter V or C inside shows whether domains are more like Ig V or C regions. The S symbols show intrachain disulphide bonds and S-S interchain bonds. N-linked carbohydrate structures are shown by $\uparrow$. For the poly Ig receptor the disulphide bonds and positions for carbohydrate structures are predicted from the sequence but not established. The estimated number of amino acids in the cytoplasmic tails of each structure is as follows: Thy-1, 0; IgM heavy chain, 3; class I MHC heavy chain, 30; class II MHC α and β chains, 15 and 10 respectively (refs 3, 11–15).

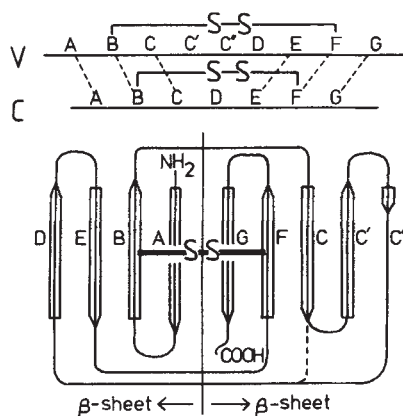


Fig. 2 β -strands in Ig V and C domains. The top part shows, as letters, the regularly spaced patches of sequence forming β -strands along the V and C domains and the position of the conserved disulphide (-S-S-) bond; the dotted lines show β strands in which sequence homology is easily seen. The lower part shows the folding pattern of V and C domains with two β -sheets laid out and separated by the line down the middle. For a C domain, β -strand C continues directly to D (see dotted line) while a V domain contains an extra loop of sequence as shown by C' and C'' (modified from ref. 2).

ligand-receptor system mainly in terms of two different molecular types¹⁰, the data of Mostov *et al.* raise the possibility that a single immunoglobulin-like domain may have been a sufficient starting point. Adhesion between cells expressing the same immunoglobulin-like molecule may have been mediated by domain-domain interactions in a manner similar to that seen, for example, between heavy-chain C domains as shown in Fig. 1.

In next week's *Nature*, data will be published suggesting that T-lymphocyte receptors are also immunoglobulin-related, a finding which will fill a major gap with regard to the molecular nature of the antigen-recognition molecules. In the wider sense it will be of interest to see whether all immunoglobulin receptors are related to immunoglobulin and to the poly Ig receptor, and whether other immunoglobulin-like molecules that function in various cell interactions or receptor functions remain to be discovered. Given the current interest in structural studies on cell-surface molecules these questions should soon be answered. □

Alan F. Williams is Director of the MRC Cellular Immunology Unit, Sir William Dunn School of Pathology, Oxford OX1 3RE.

- Mostov, K.E., Friedlander, M. & Blobel, G. *Nature* **308**, 37 (1984).
- Amzel, L.M. & Poljak, R.J. *A. Rev. Biochem.* **48**, 861 (1979).
- Kehry, M. *et al. Cell* **21**, 393 (1980).
- Edelman, G.M. *et al. Proc. natn. Acad. Sci. U.S.A.* **63**, 79 (1969).
- Richardson, J.S. *et al. J. molec. Biol.* **102**, 221 (1976).
- Steinman, H.M. *J. biol. Chem.* **255**, 6758 (1980).
- Kelin, J. *et al. Nature* **291**, 455 (1981).
- Brickell, P.M. *et al. Nature* **306**, 756 (1983).
- Williams, A.F. *J. theor. Biol.* **98**, 221 (1982).
- Jensinius, J.C. & Williams, A.F. *Nature* **300**, 583 (1982).
- Ploegh, H.L. *et al. Cell* **24**, 287 (1981).
- Kimball, E.S. & Coligan, J.E. *Contemp. Topics molec. Immun.* **9**, 1 (1983).
- Lee, J.S. *et al. Nature* **299**, 750 (1982).
- Larhammar, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3687 (1982).
- Williams, A.F. & Gagnon, J. *Science* **216**, 696 (1982).

Cosmology

From quark to cosmos

from John D. Barrow and Joseph Silk

THE early Universe has provided particle physicists with an unrivalled accelerator of high-energy particles while various elementary particles predicted by exotic theories of unification are, to astronomers, increasingly attractive candidates for the 'missing' matter in the Universe. The meeting* arranged jointly by the European Southern Observatory and CERN in Geneva was thus a timely forum for particle physicists and cosmologists to exchange ideas, especially because the discovery of the W and Z bosons with masses as predicted by the unified theory of the electromagnetic and weak nuclear forces proposed by Glashow, Salam and Weinberg has renewed confidence in the unified gauge theory route into the elementary particle world.

The simplest grand unification scheme, based on the gauge symmetry SU(5), predicts that protons should decay with a half life of at most 10^{31} yr. E. Fiorini (University of Milan) summarized the status of the continuing search for proton decay. To isolate the rare events due to spontaneous decay of protons, extensive shielding from atmospheric muons produced by cosmic ray showers is required, but, even so, the ultimate sensitivity of these experiments is likely to be limited to about 10^{33} yr by the background neutrino flux, which is high enough to produce an event $\bar{\nu}_e + p \rightarrow e^+ + n + \pi^0$, mimicking proton decay with a 10^{33} yr lifetime.

Preliminary results were reported at Geneva from experiments carried out deep underground in the Kolar goldfield, Kamioka, the Mont Blanc tunnel, the Silver King mine in Utah and the Morton Salt mine in Ohio. The last experiment provides the most sensitive limit so far, a minimum half life of 1.5×10^{32} yr for the proton. Six candidate events, not due to externally induced reactions, have been found in the other experiments but it is too soon to draw from them definite conclusions about the actual proton lifetime. But the simplest SU(5) scheme of grand unification can now be ruled out. More elaborate theories, predicting proton lifetimes of 10^{32} yr or longer, are readily constructed, however, and have various implications for astrophysics.

One success to report is the possibility of explaining the ratio of protons to photons in the Universe. Since the photons now seen in the 3K background radiation are the remnants of equal numbers of particles and antiparticles created during the thermal equilibrium of the first instants of the Universe, the observed protons represent an excess of matter over antimatter. This

asymmetry could have arisen naturally during the first 10^{-35} second (whereafter the grand unification symmetry was broken) provided that protons decayed and charge parity (CP) violation occurred. CP violation has indeed been measured in Kaon decays, but no connection with the early Universe has been demonstrated.

Another expectation of all but the simplest of SU(5) theories and of more exotic unification schemes is that neutrinos should have non-zero mass, although, as the experimenters indicated at Geneva, the theories have not been able to make useful predictions of it. Astronomers have enthusiastically welcomed this development, for even an electron neutrino mass of 10^{-6} eV could, by means of its associated oscillation into other neutrino types, help account for the observed paucity of solar neutrinos. A mass in excess of 1 eV would be significant, since neutrinos would then contribute more than stars to the mass density of the Universe. The Universe would be closed if the neutrino mass were between 25 and 100 eV, the uncertainty arising because of a factor of 2 uncertainty in the Hubble constant, whose square determines the critical closure density. The number of neutrinos is predicted precisely by the big bang theory, although the mass could be spread among several species. Several speakers (P. Darriulat, D. Schramm) pointed out that the existing CERN UA1/2 data on Z^0 decays already allows an upper bound of about twenty to be placed on the number of neutrino types, although this limit involves assuming a top quark mass of about 30 GeV.

Mössbauer (Munich) reviewed the results of neutrino mass experiments. The only positive result that is currently not contradicted by other experiments is that of Lyubimov and co-workers in the Soviet Union, which measures the energy distribution of electrons from tritium decay — energy conservation yielding information about the neutrino mass. The experiment is difficult, but the most recent result after several years of data accumulation is that the electron neutrino has a mass of at least 20 eV. Unfortunately, no members of the Soviet group were present to give a detailed account of this crucial experiment.

Of the two other classes of experiment, one measures neutrino oscillations, the conversion of electron neutrinos into mu or tau neutrinos which occurs with unknown probability if either neutrino has a finite mass. Experimental results using targets near nuclear reactors now rule out a mass difference between the two neutrino species of less than about 0.1 eV — provided the mixing probability is not very small. Unfortunately, theorists have no

*The meeting on 'The large-scale structure of the Universe' took place on 21–25 November 1983.

clue as to the likely range of mixing probability, so these experiments are not really conclusive.

If neutrinos are their own antiparticles (that is, are Majorana rather than Dirac particles), the probability of neutrinoless double β decay is greatly enhanced, and T. Kirsten (University of Heidelberg) reported a limit of 6 eV for the electron neutrino mass from studies of the decay of the geochemical abundances of ^{128}Te and ^{130}Te . Similarly, laboratory experiments on ^{76}Ge yield a limit of 14 eV. These limits conflict with Lyubimov's experiment only if the electron neutrino is of Majorana type. Direct limits on the masses of the mu and tau neutrinos are much weaker, 0.5 and 250 MeV respectively, and the tau neutrino has yet to be discovered.

At Geneva, these results set the stage for the cosmologists. Speculations are rife about the nature of the dark matter in the Universe, the case for whose pervasiveness was eloquently argued by S. Faber (University of California, Santa Cruz) on the basis of direct determination of the ratios of mass to luminosity in single galaxies, and in groups and clusters of galaxies. [A. Sandage (Mount Wilson and Las Capamas Observatories, Pasadena) showed how far we still are from a direct determination of whether the Universe could actually have a closure density in dark matter.] The candidates canvassed range from elementary particles of mass 10^{-5} eV up to super-massive stars and black holes at $10^6 M_{\odot}$, allowing some 77 decades of mass for speculation.

Neutrinos with non-zero mass provide at least a verifiable hypothesis, and theories of the large-scale structure of the Universe have now so advanced that observations of the galaxy distribution can be used to test it. The argument went like this.

Inflationary scenarios for the early expansion predict that the mean density is almost exactly at the closure value. Since the primordial synthesis of deuterium and helium in the big bang would conflict with observational evidence if the density in baryons were more than 10 per cent of the closure value (J. Audouze, Institut d'Astrophysique, Paris), the Universe can be at closure density only if non-baryonic matter is sufficiently abundant. Massive neutrinos probably cannot play this part, for reasons that have only emerged in the course of detailed comparisons with the observed large-scale structure of the Universe. Oort (University of Leiden), in an account of the very large-scale structure of the Universe, described the enormous structures, extending up to 50 Mpc, that appear in three-dimensional catalogues of the galaxy distribution. Yet, according to D. Wilkinson (Princeton University), the Universe is isotropic and homogeneous to better than 1 part in 10^4 in the distribution of the 3K background radiation when the 180° anisotropy due to our motion is subtracted. Earlier studies nevertheless suggested rather promising agree-

ment in a universe dominated by neutrinos of non-zero mass; a natural scale emerged, determined by the maximum distance neutrinos could stream freely as the universe expanded before they slowed down on account of their mass. Below this scale, roughly matching that of a supercluster of galaxies, no pre-existing fluctuations would have survived, while the first structures to collapse and form galaxies would have had supercluster scale. Ya. B. Zeldovich first showed that highly flattened pancake-like structures would inevitably develop, and fragment into galaxies.

Numerical simulations of non-linear collapse by C. Frenk (University of Sussex) and his co-workers have now revealed a possibly fatal flaw in the neutrino pancake model. The larger the first non-linear scale, the later collapse must occur if the galaxy distribution is to be clustered in the mean to the extent observed. So galaxy formation must have occurred at an unacceptably recent epoch if the Universe is at critical density due to massive neutrinos. Moreover, constraints from the small-scale isotropy of the cosmic background radiation also eliminate any plausible model for a baryon-dominated universe (J. Silk, University of California, Berkeley).

These arguments are not yet ironclad, but are sufficiently persuasive to direct cosmologists towards other candidates for the dark matter, of which the most attractive is a class of particles collectively labelled 'cold relics'. Their velocity dispersion has always been negligible over all scales of interest for large-scale structure, which implies that primordial fluctuations, arising for example in the inflationary era, would survive and eventually be gravitationally enhanced. The persistence of small-scale structure, however, implies that galaxies are the first objects to form, but there is expected to be sufficient power on large scales that pancake-like structures should still develop. Inflationary models of the Universe provide the theoretical backing for this expectation: not only can they yield a source of fluctuations but, by producing them on all scales, lead to a unique prediction for large-scale structure.

The candidates for the cold particles include massive photinos and axions. A photino is predicted by supersymmetry to be the fermionic partner of the photon, and is likely to be long-lived. The axion is a coherent field invoked in the theory of strong interactions to circumvent the prediction of a strong CP violation that is in disagreement with experimental limits. In ordinary grand unification schemes, the axion is a boson that, because of its initial coherence, is born essentially at rest when the electroweak symmetry is broken.

Despite their exotic origins, there have been recent proposals to measure the masses of both photinos and axions. Thus D. Sciamia (Universities of Oxford and Trieste) pointed out that photinos of mass

in excess of about 0.5 GeV would undergo annihilation with antiphotinos, and that while surviving photinos could conceivably be numerous enough to give a critical density, mutual annihilations would result in a possible detectable background flux of γ radiation. As for axions, while the individual mass is only about 10^{-4} eV, there would be about 10^5 axions per photon in the background radiation at closure density, and this large cosmic flux would be further enhanced by axion infall into our local galactic environment. So, in a low-temperature laboratory experiment proposed by P. Sikivie, axions would interact with a strong inhomogeneous magnetic field; those with a mass in the range required to supply the dark material in the halos of spiral galaxies, having a Compton wavelength in the microwave range, might be observable as microwave emission.

Inflationary theories as such did not receive as much attention at Geneva as at previous gatherings of particle physicists and cosmologists. Despite their appealing simplicity — no surviving magnetic monopoles, the present density within one part in a million of the closure value, large-scale homogeneity with a scale-free spectrum of inhomogeneities superimposed — theoretical justification has proved elusive. D. Nanopoulos (CERN) gave an overview of the inflationary universe model and stressed some of the severe problems that have been recently pointed out. It appears that, at present, inflation remains an appealing idea but that there is no working model that agrees with all the observational constraints.

A highlight of the meeting was the talk by S. Hawking (University of Cambridge) in which he reported recent work performed in collaboration with J. Hartle (University of California, Santa Barbara). It was shown how the quantum state of a spatially closed universe may be described by a wave function obeying a time-independent zero-energy Schrödinger equation. A simple model of an isotropic and homogeneous universe containing a massive scalar field, introduced to represent matter fields in the universe, reveals a wave function for the universe which can be represented as a superposition of different quantum states peaked in probability around particular classical lorentzian solutions of the Einstein equations. These classical solutions are non-singular, begin with a period of exponential expansion and then become matter-dominated before reaching maximum volume and recollapsing. This picture leads to a description of the most probable state for the universe which corresponds, in the classical limit, to one that is singularity-free. □

John D. Barrow is at the Astronomy Centre, University of Sussex, Brighton BN1 9QH, and Joseph Silk is at the Astronomy Department, University of California, Berkeley, California 94720.

Earth science

Dynamics of a stratified mantle

from Raymond Jeanloz

EVIDENCE is mounting that the Earth's mantle may be stratified; that is, that the upper and lower mantle are separately convecting regions. The presence or absence of mantle stratification has profound consequences for the thermal evolution of the Earth, and for the inferred pattern of tectonic movements deep beneath the surface. One purely geometrical characteristic of stratification, for example, is that mantle flow is confined within a region of large aspect ratio: the upper mantle is a mere 700 km deep, but lithospheric plates are of the order of 5,000 km wide. How could a flow pattern be maintained in the upper mantle, and what might be the geodynamic consequences of stratification? In a recent article, Greg Houseman addresses some of these questions and concludes that convection in the upper mantle can be modelled quite naturally as a large aspect-ratio flow that is predominantly driven by a mean horizontal temperature gradient¹.

A pattern of equant convection cells, with aspect ratio of about one, has been produced by a number of fluid dynamical calculations on mantle flow. Some workers have thus concluded that convection must extend through the whole mantle, without stratification. Houseman points out, however, that the nature of the flow pattern is sensitive to the boundary conditions imposed, and the appropriate boundary conditions defining flow within the Earth are not well known: convection in a stratified mantle is by no means ruled out. The presence of moving lithospheric plates at the surface, for example, could be a major factor controlling the flow in large aspect-ratio cells in the upper mantle (presuming, of course, that plates are mechanically coherent).

Another factor, which is emphasized in Houseman's work, is the horizontal temperature difference between the mantle underlying ridges (hot) and subduction zones (cold). Houseman explores the effects of varying the Peclet and Rayleigh numbers on convection in a large aspect-ratio region simulating the upper mantle: in his formulation, these parameters are varied by changing the velocity of the upper surface and the heat flux from below respectively. He finds that the horizontal temperature variation can drive the flow quite effectively, with only a 250°C average temperature difference across a large convection cell — of the same scale as would occur under the Pacific plate.

In order to determine which is the more important driving factor, the moving upper boundary or the horizontal temperature gradient, Houseman calculates the surface topography and gravity anomalies associated with his models. He

concludes that observations made at the Earth's surface can only be matched by his models if the horizontal temperature variation is the dominant driving force in the upper mantle. The results are sensitive to the parameter values that are estimated (for example, mantle viscosity), and indeed Houseman's calculations are representative of only a particular choice of boundary conditions which are not wholly realistic for the Earth's mantle (for example, assumptions of constant viscosity and zero velocity at the lower boundary; and by necessity, the use of two-dimensional computation). Nevertheless, systematic studies of convection in the upper mantle, such as Houseman's work, are gaining importance due to the increasing evidence that the mantle is stratified.

What is the evidence? First, the observed density and compressibility of the lower mantle are not well matched by values that are calculated for a mineral assemblage with the average composition of the upper mantle^{2,3}. In contrast, new elasticity data⁴ reinforce the conclusion that the upper mantle and transition zone make up a simple unstratified region to first order (if one ignores lateral heterogeneity, in particular). Further evidence is provided

by the fact that the 670-km discontinuity, which marks the top of the lower mantle, is much too sharp to be explained in terms of known phase transformations⁵: a change of composition, and hence stratification, is apparently required. Also, there is the classical observation that subduction-zone earthquakes end by 700 km depth, thus suggesting that downgoing slabs penetrate through the upper mantle but not into the lower mantle (for example ref. 6, although see ref. 7). These results are in accord with geochemical models, based on isotopic, trace element and rare gas measurements, in which the upper mantle is viewed as a depleted reservoir that has been distinct from the undepleted lower mantle for billions of years (for example, refs 8–10). Thus, studies of the large aspect-ratio flow patterns expected to occur in a stratified mantle are likely to be of crucial importance for understanding the dynamics of the upper mantle. □

Raymond Jeanloz is at the Department of Geology and Geophysics, University of California, Berkeley, California 94720.

1. Houseman, G. *Geophys. J. R. astr. Soc.* **75**, 309 (1983).
2. Anderson, D.L. *Geophys. Res. Lett.* **3**, 347 (1976).
3. Jeanloz, R. & Thompson, A.B. *Rev. Geophys. Space Phys.* **21**, 51 (1983).
4. Weidner, D.J., Sawamoto, H., Sasaki, S. & Kumazawa, M. *J. geophys. Res.* (in the press).
5. Lees, A.C., Bukowski, M.S.T. & Jeanloz, R. *J. geophys. Res.* **88**, 8145 (1983).
6. Richter, F.M. *J. geophys. Res.* **84**, 6783 (1979).
7. Rubie, D.C. *Nature* (in the press).
8. Jacobsen, S.B. & Wasserburg, G.J. *J. geophys. Res.* **84**, 7411 (1979).
9. De Paolo, D.J. *Geochim. cosmochim. Acta* **44**, 1185 (1980).
10. O'Nions, R.K. & Oxburgh, E.R. *Nature* **306**, 429 (1983).

Genetics

DNA hybridization — beware

from Roger Patient

Two papers recently published in *Nature*^{1,2} described the apparent detection of 'extra' major histocompatibility (HLA class I) antigens on the surface of human T-cell leukaemia/lymphoma virus (HTLV)-infected cells and suggested that the viral envelope (*env*) protein was similar enough in structure to HLA class I proteins to have been responsible. The latter claim was based on nucleic acid hybridization data and has now been questioned by Jennings and Even³ in the light of subsequent sequence comparisons (see p.85 of this issue of *Nature*). One is reminded of the hybridization-versus-sequence conflict described by Beverley Griffin in her *News and Views* article of 18 months ago⁴, but there are a number of differences.

Clarke *et al.*² probed HTLV, cloned in bacteriophage λ , with an HLA gene, cloned as its cDNA, that only differed from HLA-B7 by 1 amino acid out of the 48 determined by DNA sequencing⁵. The only hybridizing region in the HTLV clone was localized to the *env* gene, which hybridized to a region spanning the first three exons of HLA, enclosing the extracellular domains

of the protein. From that it was concluded that the HTLV envelope protein exhibits homology to the extracellular domains of HLA-B7; that herein lay a possible explanation for the 'extra' HLA class I specificities detected in HTLV-infected cells; and that such a masquerade on the part of HTLV would aid in the evasion of host immune detection and facilitate tumorigenesis.

After Clarke *et al.* had submitted their manuscript, the complete sequence of HTLV was published⁶. Jennings and Even searched the sequence for amino acid homologies with both HLA-B7 and the closely related HLA-A2 but found none, thus calling into question the conclusions based on the nucleic acid hybridization data of Clarke *et al.*, who in reply raise the formal possibility that the antigenic determinant might be sufficiently small to be missed by amino acid sequence comparisons. But would it then be sufficiently large to be picked up by nucleic acid hybridization? Are we faced with nucleotide homology that is not reflected at the protein level, and if so what does this

mean? Only protein homologies can have consequences at the cell surface.

It seems that we have been misled again by nucleic acid hybridization data. What is the basis of the hybridization observed by Clarke *et al.*? Clearly we need to compare the nucleotide sequences of the actual stretches of DNA which hybridize but these are so far unavailable. However, both Clarke *et al.* and Jennings and Even have compared the published HTLV nucleotide sequence⁶ with that of a class I HLA gene which exhibits homologies to both HLA-B7 and HLA-A2 (ref. 7), and no significant homology was found. The examples of deception by nucleic acid hybridization data described previously by Griffin⁴ involved a failure to detect by hybridization the significant homologies later revealed by DNA sequencing. Here we have the reverse problem: a failure to confirm, with sequence data, homology suggested by hybridization studies.

We are left with the uncomfortable business of casting around for possible explanations for the hybridization data. The sequenced HTLV clone differs from the one used in the hybridization by a number of restriction sites in the *env* region and the sequenced HLA gene differs from HLA-B7 (and thus the cDNA cloned probe) by quite a number of residues. Therefore precise comparison might yet resolve the problem. If so, would it mean that different HTLV isolates mimic different HLA class I antigens?

Beyond this we are into the realm of artefacts. The first which springs to mind is that caused by the homopolymer tails created in the construction of the HLA cDNA clone⁵. Precedent for such deception comes from *in situ* hybridization to the lampbrush chromosomes of the newt, *Triturus*, where apparent hybridization to globin cDNA clones turned out to be hybridization of poly(rG) and poly(rC) transcripts to G-C tails⁸. Another example, this time involving A-T tails, was exposed in the human β -globin locus where a region of hybridization thought to contain a pseudogene was found to be competed out by an excess of unlabelled poly(rA)⁹.

In the case under discussion, the tails are apparently oligo(G).oligo(C) and the hybridizing region of HTLV is G-C rich. When the HLA cDNA clone was split into a 5' probe and a 3' probe, both containing a G-C tail, the HTLV *env* region hybridized strongly with the 5' probe and only very weakly with the 3' probe. However, until

we know the relative lengths of the two tails, this result does not rule out artefactual hybridization due to runs of Gs and Cs. Similarly, Clarke *et al.* point out that other G-C-rich regions in HTLV do not hybridize to the cDNA probe and claim that this suggests artefactual hybridization because of G-C richness is not occurring. Unfortunately, it depends on the sequence organization. A G-C-rich fragment containing runs of Gs and Cs will hybridize whereas an equally G-C-rich fragment without such runs will not.

Astronomy

Comets colliding with the Sun

from David W. Hughes

WHY do so many comets hit the Sun? Observations of comets colliding with the Sun have come from Solwind, a coronagraph on board the US Air Force Space Test Program satellite P78-1. It records images of a 2.5–10.0 $R_{\odot}$ Sun-centred annular field with a sensitivity adequate to record features having brightnesses greater than 10^{-10} that of the solar photosphere. The satellite orbits Earth every 97 minutes and is sunlit for one hour of that time, during which it takes six coronal images.

Between 1979 and 1982 the Solwind experimenters^{1,2} discovered three comets, all of which collided with the Sun. Images of the corona taken on 30–31 August 1979 showed comet 1979 XI approach the Sun from the south-west, pass behind the coronagraph's solar occulting disc (radius 2.5 $R_{\odot}$) and not re-appear; subsequently there was a brightening of the corona. The three comets 1979 XI, 1981 I and 1981 XIII have perihelion distances of 0.35 $R_{\odot}$, 1.05 $R_{\odot}$ and 0.92 $R_{\odot}$ respectively. They are all members of a family of sungrazing (and colliding) comets discussed in some detail by H. Kreutz at the end of the last century. They are retrograde (inclination 142°) and have aphelia at about 180 AU.

Comets hitting the Sun are unusual and Paul R. Weissman (Jet Propulsion Laboratory, Pasadena) has now grappled with the celestial mechanical problem of explaining how the perihelion distances of comets can be perturbed to values less than the solar radius³.

The Oort cloud, that collection of about 2×10^{12} long-period comets spread through a 50,000-AU-radius spherical volume centred on the Sun, is being gently perturbed by passing stars. The aphelion velocity vectors of the comets are randomly scattered in velocity phase space, so after perturbation the comets are uniformly distributed as far as perihelion distance is concerned. Thus if a comet at aphelion (assumed to be 50,000 AU) has a velocity of less than 5.7 cm s⁻¹, it will collide with the Sun when at perihelion. Everhart⁴ found that 15.7 comets per AU of perihelion distance brighter than +11 absolute

In conclusion, one feels that since the hybridizing regions are small and could be sequenced very easily, this should be done as soon as possible to resolve these questions. Misleading hybridization data should not discourage us from using the technique but knowledge of the pitfalls should make us more cautious in interpreting the results. □

Roger Patient is at the Department of Biophysics, Kings' College, University of London, 26–29 Drury Lane, London WC2B 5RL.

magnitude enter the inner Solar System per year from the Oort cloud. One should hit the Sun every 68.3 years. Unfortunately, those on Sun-impacting trajectories are difficult to see, Everhart calculating that there is a probability of less than 7 per cent of discovering the comet before impact; so that the average Earth-based observer should see one every 1,000 years or so.

Comets of the Kreutz group, however, have a semi-major axis of 90 AU and are nowhere near the Oort cloud. Their inclination of 142° also keeps them away from perturbing planets. So another explanation must be sought for their small perihelion distances. And we have seen ten in the past 100 years.

Weissman considers the non-gravitational effect produced by the time lag in the emission of sublimated gases from a spinning cometary nucleus. This can decrease the semi-major axis but has only a very small effect on the perihelion distance. The calculations yield a value of 10^{-5} AU per orbit; a random walk in perihelion distance down to the photosphere grazing condition is possible for 1981 I and 1981 XIII, but it is far too slow to explain the 0.35 $R_{\odot}$ perihelion distance of 1979 XI.

May the disruption or splitting of the cometary nucleus give a fragment an orbit with such a low perihelion distance? Tidal splitting might have occurred when the comet last passed within the Roche limit of the Sun or a planet. Cometary nuclei are thought to be extremely fragile. Comet 1889 V split after passing through the Roche limit of Jupiter in 1886. And the Sun's Roche limit is 2.24 $R_{\odot}$ from its centre so all the Kreutz family have passed within it. Sekanina⁵ has shown that the typical separation velocity for the fragments is of the order of a few metres per second and this is again insufficient to produce the desired change.

Comets also break up in other places, seemingly at random. This process could be associated with heating and cooling of the nuclear surface and interior as the comet moves towards and away from the Sun. The probability of disruption

1. Mann, D.L. *et al.* *Nature* **305**, 58 (1983).
2. Clarke, M.F., Gelmann, E.P. & Reitz, M.S. *Nature* **305**, 60 (1983).
3. Jennings, P.A. & Even, J. *Nature* **308**, 85 (1984).
4. Griffin, B. *Nature* **297**, 77 (1982).
5. Sood, A.K., Pereira, D. & Weissman, S.M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 616 (1981).
6. Seiki, M., Hattori, S., Hirayama, Y. & Yoshida, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3618 (1983).
7. Malissen, M., Malissen, B. & Jordan, B.R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 893 (1982).
8. Callan, H.G. & Old, R.W. *J. Cell Sci.* **41**, 115 (1980).
9. Shen, S.-H. & Smithies, O. *Nucleic Acids Res.* **10**, 7809 (1982).

increases as the comet approaches the Sun but still the observed separation velocities are insufficient to produce the perturbation required to give the 0.00164 AU perihelion distance of 1979 XI.

Lastly, comet-comet collisions have been considered. This mechanism could change a comet's orbit or break the comet up, giving the fragments a range of velocities. The high inclination of the Kreutz orbit makes it very unlikely that they would hit an asteroid. And Weissman⁶ calculates that they have a probability of only 2.8×10^{-16} per orbit of hitting a long-period comet. But even though this process seems unlikely, it remains at the moment the most scientifically plausible explanation for the peculiar orbit of 1979 XI.

So 1979 XI remains a mystery and there seem to be only two obvious possibilities. Either it has suffered a collision in space — an event which has an extremely low probability — or the orbit has been calculated incorrectly. □

David W. Hughes is Lecturer in Physics and Astronomy at the University of Sheffield, Sheffield S3 7RH.

1. Michels, D.J., Sheeley, N.R., Howard, R.A. & Koomen, M.J. *Science* **215**, 1097 (1982).
2. Sheeley, N.R., Howard, R.A., Koomen, M.J. & Michels, D.J. *Nature* **300**, 239 (1982).
3. Weissman, P.R. *Icarus* **55**, 448 (1983).
4. Everhart, E. *Astr. J.* **72**, 1002 (1967).
5. Sekanina, Z. *Icarus* **33**, 173 (1978).
6. Weissman, P.R. *Astr. Astrophys.* **118**, 90 (1983).



100 years ago

MR. BURNHAM'S DOUBLE-STAR MEASURES

THE recently published volume of the *Memoirs of the Royal Astronomical Society* contains a further series of measures of double stars by Mr. S. W. Burnham, made with the 18-inch refractor of the Observatory at Chicago. This series comprises measures of 151 double stars discovered by this eminent observer, which brings up the number of such objects discovered by him during the last ten years to no fewer than 1013, amongst which are included some of the most interesting stars of this class; also measures of a selected list of double stars, 770 in number, made chiefly in the years 1879 and 1880, with an appendix, the results of observations of several objects, as late as the middle of the past year.

Among the more noteworthy stars included in Mr Burnham's new Catalogue, the following may be mentioned:-

1. 126 Tauri (β 1007), "a most remarkably close and difficult pair, one of the closest known"; magnitudes 6.0 and 6.2. With a power of 1400 there was only a slight elongation.
2. B.A.C. 346; Mr Burnham thinks the principal star may be variable.
3. β Delphini (β 151). — A very rapid binary; since its detection by Mr Burnham in 1873, there has been an increase in the angle of about 180° , and a diminution in distance from 0".6 to 0".25.

From *Nature* **29**, 409, 28 February 1884.

Endocrinology

New model for steroid hormone receptors?

from William T. Schrader

ONE aspect of steroid hormone receptor action has been elevated to the status of dogma since its elucidation fifteen years ago. Following the first detection of oestradiol receptor in rat uterine cytosol¹, E. Jensen and his collaborators at the University of Chicago demonstrated by cell fractionation and autoradiography that within an hour of hormone administration, the receptor accumulates in the nucleus of target cells². Their findings led to the now-familiar 'two-step' hypothesis: oestrogen receptors are cytoplasmic proteins which, upon binding of the hormone, acquire an affinity for the nucleus to which they move. The methods were repeated and extended to numerous other steroid hormone receptor systems with similar conclusions. Now two papers^{3,4} in *Nature* challenge the dogma: it seems that steroid receptors may not be cytoplasmic proteins after all.

Over the years, there have appeared a few reports suggesting that the two-step model may not be accurate. Thus Pietras and Szego⁵ showed that hypotonic buffers used prevalently in the field for cell homogenization yielded receptors in the cytosol, while use of 0.25M sucrose resulted in localization of unoccupied receptor predominantly in the particulate fraction. Similarly, Linkie and Siiteri⁶ found that 'transformation' of the receptor from 4S to 5S forms, obligatory in the nuclear uptake idea, might be an intranuclear event; and oestrogen receptor of rooster liver was detected exclusively in the nucleus, regardless of the hormonal status⁷. More recently, Walters *et al.*⁸ measured vitamin D₃ receptor localization as a function of salt concentration in the extraction buffer and observed a pronounced effect of the ligand upon the salt molarity at which the receptor partitioned into the cytoplasmic fraction. In general, these isolated reports have been dismissed as either special cases or as insufficiently convincing evidence to warrant opening the arguments of the previous decade.

The reasons for the reluctance to continue the debate were not entirely due to intransigence on the part of workers in the field. Rather, the studies invariably used cell homogenization and its attendant artefacts as part of the experimental approach and no viable alternative strategy had emerged. Furthermore, the field had moved on to other pressing issues, such as receptor protein structure and interactions with genes.

Now the issue has been raised anew, in the pair of papers appearing in *Nature* last week^{3,4}. The two approaches differ in methodology but reach the same conclu-

sion: apparently the unoccupied oestrogen receptor is not a cytoplasmic protein in intact cells. The first method used monoclonal antibodies to human oestrogen receptor. W. King and G. Greene³ at the University of Chicago performed immunocytochemical assays of oestrogen receptor in cell cultures and in reproductive tract. Immunoreactivity was observed exclusively in the nucleus, quite inconsistent with the distribution predicted by classical biochemical analysis. The method does not, however, exclude the possibility that some unoccupied receptor may be cytoplasmic, but merely invisible because of a diffuse distribution in that compartment; or that receptor redistribution may have taken place during cell permeabilization.

The accompanying paper⁴, from J. Gorski's laboratory at the University of Wisconsin-Madison, tends to clarify these questions. Gorski and his colleagues exposed pituitary tumour cells to cytochalasin, and then prepared enucleated 'cytoplasts' by centrifugation of the treated cells. In the absence of hormone, the classical view would predict that oestrogen receptors remain in the cytoplasmic fraction. In fact, what was seen was a 10-fold reduction in receptor titre in the enucleated cytoplasts compared with whole cells. The missing receptor was recovered in the nuclear, or karyoplast, fraction.

Since the two approaches employ methodology quite different from earlier studies, it seems that a fresh round of controversy is bound to start. The original reports from the Chicago group using oestrogen receptor antibody immunocytochemistry employed different fixation methods, and showed perinuclear staining⁹. Is the difference from the current work because of fixation artefacts, choice of antibody, or what? One is always perplexed when forced to choose between two equally convincing sets of micrographs.

There remain numerous problems to be addressed before the dust will settle again. First, are the receptors described in the two recent papers really in the nucleus, or merely on the nucleus? Since active hormone-regulated genes seem to be located in the nuclear matrix¹¹ future work will need to allow detection of receptors associated specifically with active genes. Unfortunately, that is technically difficult. Second, are receptors for all steroids similarly distributed? Particularly important in this regard will be the findings for glucocorticoids, which have provided the most con-

vincing evidence so far for two steps¹². Finally, if steroid receptors are indeed weakly held non-histone nuclear proteins, what are the implications for the 'two-step' hypothesis? Clearly, the greatest damage will be to the countless slides and book diagrams of steroid hormone action which will need to be replaced. The potential economic benefits to laboratory illustrators and photographers are thus not inconsequential. However, to the rest of the world the change may be more cosmetic than substantive. It remains undeniable that hormone agonists do alter gene expression, and hence some aspect of the receptor structure must change by interaction with its ligand. That change, once widely believed to result in a compartmental shift, may now be even more subtle. Perhaps receptors are permanently associated with the genes they regulate, and thus have only to rotate or change conformation to achieve their effects. Seen in this context, there still are two identifiable states of a steroid receptor: an active (ligand-associated) state and an inactive (unliganded) one. The real significance of the

original two-step idea is retained; steroid-receptor complexes are nuclear regulatory elements wherever they originate in the cell.

The recent findings thus renew old problems, but at least may serve to incite us to the more difficult questions of receptor function in molecular terms. □

William T. Schrader is at the Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030.

1. Toft, D.O. & Gorski, J. *Proc. natn. Acad. Sci. U.S.A.* **5**, 1574 (1966).
2. Jensen, E.V. *et al. Proc. natn. Acad. Sci. U.S.A.* **59**, 632 (1968).
3. King, W.J. & Greene, G.L. *Nature* **307**, 745 (1984).
4. Welshons, W.V., Lieberman, M.E. & Gorski, J. *Nature* **307**, 747 (1984).
5. Pietras, R.J. & Szego, C.M. *J. Steroid Biochem.* **11**, 1471 (1979).
6. Linkie, D.M. & Siiteri, P.K. *J. Steroid Biochem.* **9**, 1071 (1978).
7. Snow, L.D. *et al. J. Steroid Biochem.* **9**, 1017 (1978).
8. Walters, M.R., Hunziker, W. & Norman, A.W. *J. biol. Chem.* **255**, 6799 (1980).
9. Jensen, E.V., Greene, G.L., Closs, L.E., DeSombre, E.R. & Nadji, M. *Rec. Prog. Horm. Res.* **38**, 1 (1982).
10. Barrack, E.R., Hawkins, E.F., Allen, S.L., Hicks, L.L. & Coffey, D.S. *Biochem. biophys. Res. Commun.* **79**, 829 (1977).
11. Ciejk, E., Tsai, M.-J. & O'Malley, B.W. *Nature* **306**, 604 (1973).
12. Munck, A. & Foley, R. *Nature* **278**, 752 (1979).

Goe and catche a falling starre

ACROSS the arcs of background stars, followed over a three-hour exposure, streaks a brilliant fireball trail. The picture below left, which brings to mind the words of John Donne in the title of this article, documents the first time that photography successfully pinpointed the site of a meteorite fall from the trail of a plummeting meteor and only the second time that it facilitated calculation of the orbit prior to Earth encounter.

A single photograph like this one, of course, provides limited information — the

magnitude of the meteor, its duration and its direction of motion. The dashed trajectory results from a chopping camera shutter which further indicates the velocity and rate of deceleration. But any hopes of recovering whatever meteoritic debris may have survived the swift passage through the atmosphere must depend on triangulation using a set of photographs from several stations.

It was for just that purpose that the sixteen stations of the Prairie Network were built in the American mid-west in

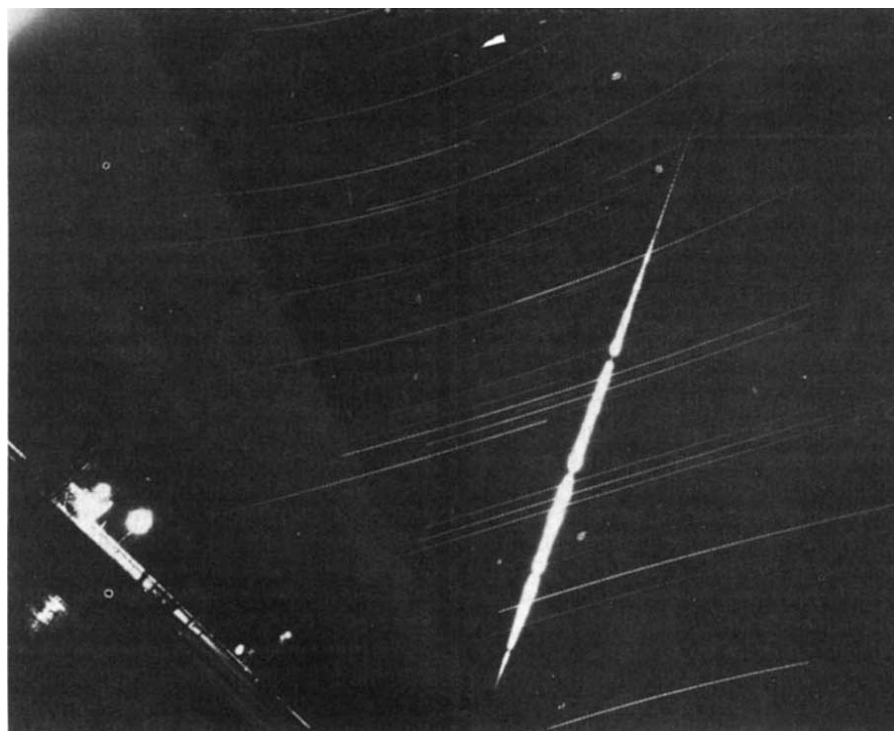
1963 by the Smithsonian Astrophysical Observatory. Each station functioned automatically, a twilight photometer triggering four aerial T-11 cameras, one directed to each cardinal point. Save for a blind spot at the zenith, the whole sky was covered at every station by the f/6.3 Planigon lenses with an 85° field of view. Plate-glass windows kept the cameras proof against rain and snow, and a cloud detector avoided waste of film. In thirteen years of scanning the skies nightly, the Prairie Network logged over 2,700 meteors (excluding bright meteor showers) but produced only one photographic recovery: the Lost City meteorite pictured below right.

The probable location of the meteorite was deduced from four photographs recorded on 3 January 1970 at two stations in Kansas and two in Oklahoma (including the one below left from Hominy). The meteor looked a good candidate for a surviving meteorite because of its low initial velocity (14.2 km s⁻¹; it can be as much as five times faster), its very low terminal velocity (3.5 km s⁻¹) and terminal height (19.5 km) and its relatively long duration (9 s).

Several factors conspire to determine the point of impact. Aside from the triangulated position and velocity inferred from the trajectory, the meteorite hunter needs to take account of the body's density and shape (to calculate the aerodynamic drag coefficient), the air density and local winds. A radiosonde flown over Oklahoma City provided the latter, but in the absence of better information one could only assume that the body itself was spherical and of normal meteoric density (3.4 g cm⁻³). The predicted fall lay a few kilometres east of Lost City, Oklahoma.

The meteorite was duly recovered from a roadside snow bank only 600 m from the target. Three more fragments were subsequently located — 17 kg altogether. Lost City turned out to be a bronzite chondrite, chemically and radiogenically dated at 4.8 to 5.5 × 10⁹ yr BP. Orbital calculations place aphelion inside the orbit of Jupiter. An asteroidal origin should not be glibly inferred, however, since both Jupiter and the Earth may have severely modified its primaevial orbit. **Jon Darius**

Selected from Beyond Vision by Jon Darius of the Science Museum, London (to be published in April by Oxford University Press). Photographs reproduced by courtesy of C.-Y. Shao, Smithsonian Astrophysical Observatory.



Poliovirus antigenic sites and vaccines

SIR — Two recent News and Views commentaries have discussed the identification of a neutralizing antibody-inducing site on poliovirus¹ and the possibility of using corresponding, synthetic peptides as vaccines². Some conclusions drawn and definitions used in these articles differ from our view of the subject matter.

In the first commentary, the author explains how recent studies^{3,4} on poliovirus type 3 (PV3) have identified an amino acid sequence of the viral structural protein VP1 which "represents the site on the virus involved in virus neutralization"¹. The implication that only a single antigenic site is involved in the induction of neutralizing antibodies is probably premature for PV-3. It is certainly incorrect for poliovirus, type 1 (PV-1)⁵⁻⁷.

For clarity, we will use the following definitions and abbreviations. The term neutralization antigenic site (N-Ag) of a virus describes that region of a capsid protein, composed of contiguous (or non-contiguous) amino acids, that induces and binds neutralizing antibodies. A neutralization epitope (N-Ep) is the molecular conformation of a sequence of contiguous (or non-contiguous) amino acids of the N-Ag that a neutralizing monoclonal antibody (N-mcAb) will recognize and bind. A single N-Ag may be capable of expressing many N-Eps. The N-Eps in such a cluster may or may not overlap functionally, that is, genetic variants of a virus resistant to neutralization by one N-mcAb may or may not escape neutralization by other N-mcAbs of the same cluster^{5,6}.

P. Minor and D. Evans and their colleagues elicited N-mcAbs against PV-3 that recognized different N-Eps. Surprisingly, all of these N-Eps belong to a single N-Ag on capsid protein VP1, as shown by genetic and biochemical methods^{3,4}. With PV-1 we have used a similar approach^{6,7}. Two N-Ags on capsid protein VP1 were identified by a binding assay (ELISA) of groups of N-mcAbs with two different virus-specific, synthetic peptides. Each group of N-mcAbs recognized a different N-Ep as shown by genetic analyses⁶. A third N-Ag on VP1 was identified by the ability of a third peptide, itself unable to bind to any N-mcAbs available to us, to "prime" a neutralizing immune response⁷. Two other N-Ags exist, one each on the VP2 and VP3 structural proteins of PV-1^{5,8,9}. In the case of VP2 the exact amino acid sequence involved is known⁹.

PV-1 therefore has at least five different N-Ags. This does not mean, however, that every animal inoculated with poliovirus will produce antibodies to all of these sites. We believe that two of the N-Ags of PV-1 are more readily recognized than the others by the animal's immune system, a phenomenon that would explain why N-mcAbs can generally be elicited only

against these two "immunodominant" sites⁷. But it should be kept in mind that the current assignment of dominant N-Ags is solely based upon results obtained with N-mcAbs. We suggest that the N-Ag identified for PV-3 (ref.3) is such an immunodominant site, but we predict that PV-3 has several other neutralization sites that are recognized to a lesser degree by the immune system.

Neutralizing antibodies may destroy the infectivity of the poliovirion by different mechanisms. The N-mcAbs directed against capsid polypeptide VP1 alter the isoelectric focusing point (pI) of PV-1⁶. Neutralization requires bivalent binding^{10,11}. This is in contrast to neutralizing antibodies directed against capsid polypeptide VP3 that do not induce a change of pI and that do not require bivalent binding for inactivation^{5,6,10}. In view of all of this, which peptides should be selected and from which regions of the capsid proteins should they originate in order to construct a synthetic vaccine?

We have tested so far 6 peptides, 5 corresponding to sequences of capsid protein VP1⁷ and one to capsid protein VP2⁹, for their ability to elicit a neutralizing antibody response. Success was scored only with a peptide corresponding to the amino acid sequence of one of the immunodominant N-Ags of VP1 and with a peptide corresponding to an amino acid sequence of VP2. Does this result imply that these peptides are the only ones "whose further potential as a peptide vaccine is worth exploring"²?

No rules have been found that can predict whether or not a peptide will mimic the N-Eps of a N-Ag. The final selection of the sequence of a peptide and its chain length must still be determined empirically. But even if a peptide carrying the sequence of a N-Ag is unable to elicit neutralizing antibodies, it can be "seen" and memorized by the immune system of the

test animal. We have shown recently that the recall of this immunological memory to a peptide can lead to a powerful "priming" effect of a neutralizing immune response⁷.

Some viruses, such as foot-and-mouth disease virus (an aphthovirus belonging to the same family as poliovirus) and influenza virus, can tolerate the rapid genetic change of their dominant N-Ags without loss of viability. This phenomenon is highly advantageous for a virus as it distracts the attention of the immune system from other, less variable N-Ags, and leads to variants which are resistant to neutralization. "Priming" of a neutralizing immune response with peptides, we speculate, may be useful in alerting the immune system not only to dominant but also less dominant (and possibly less variable) N-Ags, which, in turn, may produce long-term protection against viruses undergoing antigenic drifts or shifts.

ECKARD WIMMER

BRADFORD A. JAMESON

*Department of Microbiology,
School of Medicine,
State University of New York
at Stony Brook,
Stony Brook, New York 11794, USA*

EMILIO A. EMINI

*Division of Virus and Cell Biology,
Merck, Sharp and Dohme
Research Laboratories,
West Point, Pennsylvania 19486, USA*

1. Brown, F. *Nature* **304**, 395-396 (1983).
2. Newmark, P. *Nature* **305**, 9 (1983).
3. Minor, P.D. *et al.* *Nature* **305**, 674-679 (1983).
4. Evans, D.M.A., Minor, P.D., Schild, G.C. & Almond, J.W. *Nature* **304**, 459-462 (1983).
5. Emini, E.A., Dorner, A.J., Dorner, L.F., Jameson, B.A. & Wimmer, E. *Virology* **124**, 144-151 (1983).
6. Emini, E.A., Kao, S.-Y., Lewis, A.J., Crainic, R. & Wimmer, E. *J. Virol.* **46**, 466-474 (1983).
7. Emini, E.A., Jameson, B.A. & Wimmer, E. *Nature* **304**, 699-703 (1983).
8. van Wezel, A.L., van der Marel, P., Hazendonk, T.G., Boer-Bak, V. Henneke, M.A.C. *J. biol. Standard* (in the press).
9. Emini, E.A., Jameson, B.A. & Wimmer, E. in *Modern Approaches to Vaccines* (eds Chanock, R. & Lerner, R. (Cold Spring Harbor Press, New York, in the press).
10. Emini, E.A., Ostapchuk, P. & Wimmer, E. *J. Virol.* **48**, 547-550 (1983).
11. Icenogle, J. *et al.* *Virology* **127**, 412-425 (1983).

Availability of Mo and HTLV-II

SIR — Following recent publications, I have had numerous inquiries regarding the availability of the Mo B-lymphoblast line — an Epstein-Barr virus-transformed cell line from the patient (Mo) which is productively infected with HTLV-II — and of HTLV-II and lymphokine producing cell, derived by transforming normal T lymphocytes with HTLV-II¹. The Mo-B line and various HTLV-II-transformed T-cell lines are available upon reasonable request. So too are HTLV-II molecular probes², HTLV-I-infected primary cells (ME) and HTLV-I-infected transformants which produce lymphokines³.

I should also like to mention the position regarding the original Mo T-lymphoblast line, first isolated in 1978 and subsequently restricted because of a University of California agreement stipulating that the cell

line would not be made generally available until a patent is issued to the university. A patent on the cell line will be issued before 14 April 1984 and the cell line will then be available from the American Type Tissue Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852.

The problem of availability of biological reagents is a complex and important issue. Certainly, biological materials should be available to the scientific community for research purposes; however, the availability of these materials for commercial purposes needs to be regulated. Many factors enter into the equation of how biological reagents are made available. The "ownership" question needs to be established even if the transaction is between nonprofit organizations. Guidelines related to the exchange of reagents

should be clarified in writing with respect to how the reagents will be used, collaborative agreements, and restrictions. Many institutions, including the University of California, have generated formal agreement letters for providing biological reagents to qualified investigators at both nonprofit and for-profit institutions. These letters usually contain statements of ownership by the parent institution and include wording regarding lack of warranty and responsibility for the materials sent. They indicate that the materials are not to be used for commercial purposes nor passed to others without prior written agreement. These written agreements are time consuming but important if painful disagreement or lawsuits, such as occurred with the KG-1 cell line, are to be avoided⁴.

The question of biosafety is more important than all of the previous considerations. The history of the Mo cell line is a good example. At the time the Mo cell line was derived, we were unaware that it harboured a human T-leukaemia virus⁵. The patient, Mo, is alive and well; however, the cell line has evolved HTLV-II species that may contribute towards the high degree of malignancy of the cell line². Mo is the only individual thus far reported to harbour this rare virus. Had the cell line been made generally available before the identification of the virus, HTLV-II would have been passed to investigators all over the world with no knowledge of the bio-

hazard involved. Thus, there can be considerable danger in providing cell lines before they are completely characterized.

Lastly, there remains the burden on the originator laboratory of providing DNA, RNA, or proteins derived from these cell lines to laboratories that do not have the facilities to carry human virus-containing cells. For a small laboratory these activities can be incapacitating. The letters, phone calls, preparation of materials for shipment, and actual shipment can severely affect a research laboratory's ability to function. Some materials such as purified lymphokines cannot be made generally available because they are produced in minuscule quantities at great expense.

In summary, all biological reagents developed by investigators associated with my laboratory at UCLA are available to the scientific community at the discretion of the originating investigator and in conformity with the regulations of the University of California.

DAVID W. GOLDE

*Division of Hematology-Oncology,
UCLA School of Medicine,
Los Angeles, California 90024, USA*

1. Chen, I.S.Y., Quan, S.G. & Golde, D.W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7006-7009 (1983).
2. Chen, I.S.Y., McLaughlin, J., Gasson, J.C., Clark, S.C. & Golde, D.W. *Nature* **305**, 502-505 (1983).
3. Gasson, J.C., Chen, I.S.Y., Westbrook, C.A. & Golde, D.W. in *Normal and Neoplastic Hematopoiesis* (eds Golde, D.W. & Marks, P.A.) 129-139 (Liss, New York, 1983).
4. *Wall Street Journal*, p.1, 3 December 1980.
5. Kalyanaraman, V.S. *et al. Science* **218**, 571-573 (1982).

Limitations of physical theory

SIR — Your recent view on reducing biology to physics¹ ignores a fundamental problem. If biology is in principle reducible to physics, then so is psychology. Physics is quantum so that the prediction of behaviour would be in the form of a probability distribution of states: which state is observed is found only by observation. Since the observer has been included in the physics, there is no way to determine the outcome unless an observer external to the physics is introduced. This observer is then unexplained by the reduction.

This incompleteness of physics is a consequence of Gödel's theorem. Any comprehensive physical theory must explain structures that can model arithmetic such as human beings and computing machinery. Such a physical theory is then subject to Gödel's theorem and is necessarily incomplete and undecidable. That is, for every algorithm of the theory, there are physical events that cannot be decided by that algorithm. Thus, every comprehensive physical theory, no matter how fully developed, cannot predict the occurrence or non-occurrence of certain events. This incompleteness is a property of the theory; the outside universe may be complete but no theory can prove this fact.

This application of Gödel's theorem to physics is largely metaphysical, but may be related to the existence of the large

dimensionless numbers first noted by Dirac². This number D , is of the order of 10^{39} and is roughly equal to the age of the Universe in atomic units, to the square root of the number of protons in the Universe and to the ratio of gravitational to electromagnetic force. In general, D appears in many relations between macrophysical and microphysical quantities. Dirac postulated that D was a constant with resulting gravitational constant variation. Carr and Rees³, among others, have explained D on the basis of the anthropic principle: observers would not exist to observe this number unless the Universe was such that D had the proper value and relationships. Assuming that Gödelian incompleteness occurs for physical systems with complexity G (measured as the exponential of the entropy or of the negative likelihood), G will be a large number. A human body becomes Gödelian on timescales of the order of a second, so after about 3×10^{13} (infrared) vibrations of the 3×10^{24} molecules in the brain (assuming an average molecular weight of 300) prediction is necessarily incomplete. This crude estimate of 10^{38} for G is relatively near D , but suggests lines of research. For example, the size of the quantum of action may be a result of constructing a theory that applies at universal scales. Application to objects smaller order than $1/D$ times the Universe

must bring in incompleteness: for current theories this incompleteness is described by quantum mechanics. Present computing machines process only 100 bits at most 10^9 times per second so that they are far from becoming Gödelian on human timescales.

Any comprehensive physical theory is necessarily incomplete because of Gödel's theorem. This may not reflect incompleteness of reality, only a limitation on the construction of theories using finite sentences, but it is a fundamental limit to reductionism. Attempts to reduce complex systems to physics may succeed formally, but the prediction is a probability distribution of outcomes rather than a definite event. Determining the actual outcome requires another level of observation outside the physics in analogy to the resolution of incompleteness in mathematics by adding new axioms.

THOMAS L. CLARKE

*National Oceanic and Atmospheric
Administration,
Environmental Research Laboratories,
4301 Rickenbacker Causeway,
Miami, Florida 33149, USA*

1. Maddox, J. *Nature* **306**, 311 (1983).
2. Dirac, P.A.M. *Nature* **139**, 323 (1937).
3. Carr, B.J. & Rees, M.J. *Nature* **278**, 605-612 (1979).

Interstitial fluid

SIR — Bretag asserts¹ that for isolated mammalian tissue synthetic interstitial fluid (SIF)² is superior to Krebs' solution. In one respect it is almost certainly inferior, namely the glucose content.

The glucose content of Bretag's solution is 5.55 mM, or about half of that in Krebs'³. This is a crucial difference, because although Bretag's is very close to the glucose concentration in normal interstitial fluid, it fails to take into account that in excised tissue the glucose in the bathing fluid often has to diffuse further to its site of utilization than in the body, where the source of glucose is in a neighbouring capillary. In this situation Bretag's solution is effectively hypoglycemic.

In a study⁴ of the ambient glucose concentration required at 37°C to preserve long-term excitability of myelinated axons in excised vagus nerve of rabbit, 5 mM glucose sufficed in nerve from which the perineurium had been removed but not when the perineurium had been retained, when a concentration in excess of 10 mM was required. The transverse section of these nerves was 0.6×0.4 mm.

Bretag's formula for SIF would be improved by replacing most or all of the 7.6 mM sucrose with glucose.

B. RAYMOND FINK

*Department of Anesthesiology,
University of Washington
School of Medicine,
Seattle, Washington 98195, USA*

1. Bretag, A.H. *Nature* **307**, 116 (1984).
2. Bretag, A.H. *Life Sci.* **8**, 319-329 (1969).
3. Krebs, H.A. *Biochim. biophys. Acta* **4**, 249 (1950).
4. Fink, B.R. & Cairns, A.M. *Pain* **12**, 307 (1982).

Global atmospheric effects of massive smoke injections from a nuclear war: results from general circulation model simulations

Curt Covey*, Stephen H. Schneider & Starley L. Thompson

National Center for Atmospheric Research, Boulder, Colorado 80307, USA

We report three-dimensional calculations of regional and global climatic effects of smoke generated by a large-scale nuclear war. Tropospheric aerosols of absorption optical depth 3, when injected into Northern Hemisphere mid-latitudes and maintained for 1–3 weeks, cause intense radiative heating of the mid-troposphere with substantial surface cooling over land. Mid-latitude surface temperatures in continental interiors can drop well below freezing in a matter of days regardless of season. Our results, although based on several assumptions, suggest that circulation changes caused by aerosol-induced atmospheric radiative heating could spread the aerosols well beyond the altitude and latitude zones in which the smoke was initially generated.

LARGE fires have long been recognized as a possible consequence of a nuclear exchange¹ but with few exceptions (for example, see ref. 2) their regional and global scale environmental consequences have been ignored until recently. Crutzen and Birks³ have pointed out, via a simple order-of-magnitude estimate, that fires ignited by a full-scale nuclear war could easily produce enough smoke to block out sunlight (that is, produce optical depths of order unity) over much of the Northern Hemisphere for a period of weeks or longer. Subsequent studies^{4–7} confirm the magnitude of the problem and thus raise, for the first time, the possibility that global atmospheric consequences of nuclear war could be significant even when compared with the obviously devastating nature of the direct effects (blast, heat, short-term fallout)⁸. In this study we investigate, by using a three-dimensional general circulation model (GCM), the atmospheric consequences over a few weeks of smoke generated by a nuclear war—that is, perturbations in atmospheric and surface temperatures, winds, etc.

Turco *et al.*^{4,5} studied the climatic implications with a one-dimensional annually averaged radiative-convective model (RCM) which averages all horizontal variations and considers quantities such as temperature and aerosols to be functions only of altitude (see ref. 9). When the heat capacity of the surface was set equal to a small value characteristic of a layer of soil, smoke aerosol injection scenarios resulted in a substantial decline in surface temperature (characteristically $\sim 30^\circ\text{C}$ within 30 days, or enough to lower the global mean annual temperature well below the freezing point). At the same time the middle troposphere warms; both events are caused by the absorption of sunlight by mid-tropospheric smoke. But when Turco *et al.* used a surface heat capacity representative of an ocean mixed layer, only small temperature changes resulted ($2\text{--}3^\circ\text{C}$ after 6 months). Of course, in the real atmosphere under the impact of the aerosol loading, land areas would be expected to cool much more substantially than the oceans, and those areas of the globe not covered by smoke would suffer a much smaller temperature perturbation (though precipitation or other weather elements could still be substantially perturbed). Turco *et al.*⁴ extrapolated the results of others' simple climate models to estimate that ocean-to-land heat transport would reduce the land surface temperature drop by $\sim 20\%$ in the middle of continents and by 40% near the coasts. However, they also pointed out that the same winds that would ameliorate the

surface freezing could—when enhanced by aerosol-induced temperature contrasts—spread the aerosols well beyond their original latitudes of injection, perhaps all the way into the Southern Hemisphere.

The conclusions of Turco *et al.* have been extended by calculations with higher-dimensional models for different, yet fairly comparable scenarios of atmospheric smoke injection. MacCracken¹⁰ reports results from both a one-dimensional RCM and a two-dimensional (latitude/height) model. The one-dimensional model gave a land surface cooling of $\sim 30^\circ\text{C}$, in close agreement with Turco *et al.* The two-dimensional model, which parameterizes three-dimensional atmospheric motions to account approximately for thermal mixing, gave a $\sim 15^\circ\text{C}$ cooling of land areas underlying the smoke. A three-dimensional model—but with only two layers in the vertical—has yielded similar results¹¹.

Clearly, the issues raised above should be addressed by fully three-dimensional models which include the important radiative processes present in RCMs (and the real atmosphere), and which also account for regional and seasonal heterogeneity and can calculate explicitly large-scale winds and their advection of heat—that is, three-dimensional atmospheric general circulation models. Ideally one should use a model which also accounts for aerosol transport and removal interactively, that is, calculates aerosol spreading and removal based on the atmospheric temperature structure and circulation which results from aerosol-induced heating. We have used a GCM developed at the National Center for Atmospheric Research (NCAR)—the Community Climate Model (CCM)—in an initial study which assumes a predetermined smoke aerosol distribution. As yet we do not have the capabilities to perform a three-dimensional fully interactive radiation, transport and removal calculation. Such numerical experiments must await considerable model development and validation.

The model

The basic CCM is documented elsewhere^{12–14} so we will give only a brief description here. The model uses the spectral transform technique for horizontal discretization with rhomboidal truncation at wavenumber 15, corresponding to a horizontal resolution of about 4.5° in latitude and 7.5° in longitude. In the vertical there are nine layers which encompass the troposphere and the stratosphere. (Relatively fine vertical resolution is necessary to compute adequately the important changes in atmospheric structure and circulation which result from the heating of the smoke.) Atmospheric winds, temperature and moisture are computed using standard conservation laws

* Present address: Division of Meteorology and Physical Oceanography, Rosentiel School of Marine and Atmospheric Sciences, University of Miami, Miami, Florida 33149, USA.

together with semi-empirical parameterizations which govern phenomena too small in scale to be resolved and computed explicitly by the model.

The CCM cloud/radiation scheme, of particular interest to our proposed study, is described by Ramanathan *et al.*¹⁵. The parameterization includes interactive clouds which form and disappear as determined by relative humidity and convective activity. Absorption of sunlight within the atmosphere by ozone, water vapour, carbon dioxide and molecular oxygen, with enhanced absorption and scattering in clouds, follows Sasamori *et al.*¹⁶ and Lacis and Hansen¹⁷, with cloud albedos constrained by available data. Long-wave processes are calculated by standard algorithms using improved values for water vapour and cloud emissivities¹⁵.

In the CCM, as in most atmospheric GCMs, sea surface temperatures are prescribed as a lower boundary condition. This means that the small cooling of the sea surface expected after a few weeks of a nuclear war aerosol perturbation will not be

simulated; instead, the sea surface will remain at its prescribed temperature and will release slightly more latent and sensible heat into the atmosphere than would be the case if it were allowed to cool interactively. On the other hand, the land surface is assumed to have zero heat capacity in the CCM (land surface temperatures are calculated by assuming no net energy flux across the surface) and so the initial cooling rate of the land surface will be overestimated. Neither of these two approximations should seriously compromise the qualitative inferences from the results obtained for times of a few days and up to a month. Of course, the detailed evolution of the regional response should not be taken literally.

Simulation results

In our study, we forced the model in Northern Hemisphere mid-latitudes to absorb virtually all incident sunlight in the middle troposphere, leaving almost none to be absorbed by the surface. This forcing represents an optically thick layer of smoke which absorbs but does not scatter sunlight. For practical considerations, IR absorption and emission of the aerosols was not included. Fortunately, as IR absorptivity of smoke aerosols is about an order of magnitude less than solar absorptivity^{4,5,18}

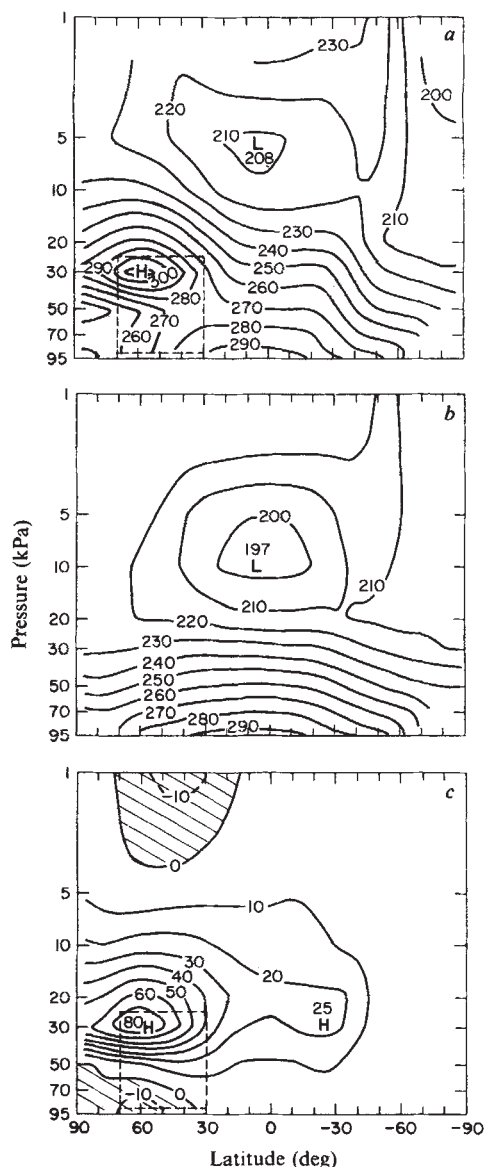


Fig. 1 Zonally averaged temperatures (Kelvin), time-averaged for $t = 10$ – 20 days after the addition of smoke to the atmosphere in the 'summer' case. The perturbed case (smoke experiment, *a*), control case (equilibrium simulation without smoke, *b*) and the difference between the two (perturbed–control, *c*) are shown. Areas with negative numbers in the difference map (cooling) are shaded. The dashed lines enclose the approximate region in which smoke was added to the atmosphere.

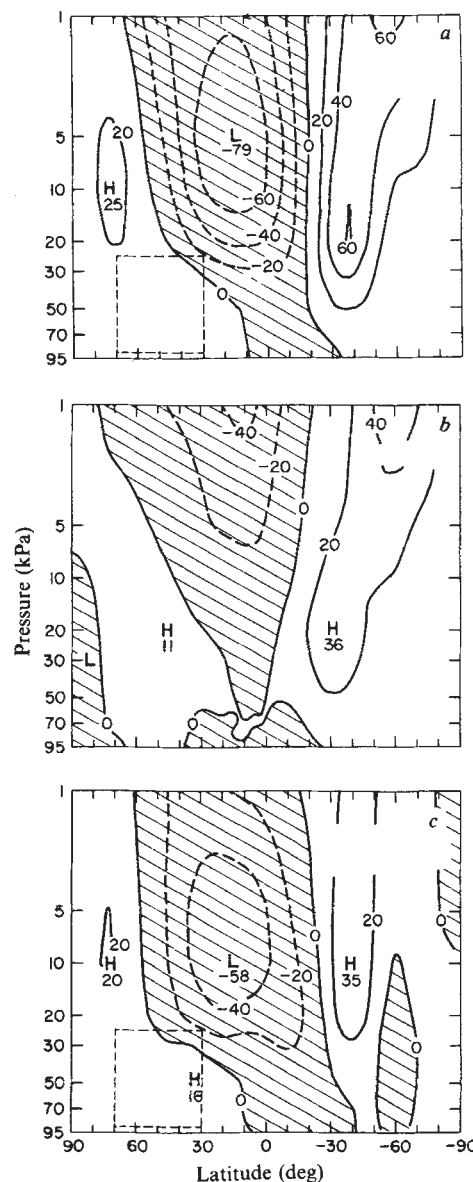


Fig. 2 Same as Fig. 1 for zonal winds (m s^{-1}) in the 'summer' case. Westerly winds are by convention positive and easterly winds are negative (shaded).

our study will delineate a first-order approximation to the problem. We used a version of the CCM which includes specified seasonal variations of solar zenith angle, ozone, ocean surface temperatures, snow and sea ice cover developed by R. M. Chervin of NCAR in order to assess the seasonality of the atmospheric effects of the smoke injections.

In making assumptions about the amount and distribution of smoke aerosols generated by a nuclear war, we followed the US National Academy of Sciences 'baseline' case, which is intended to represent conservative estimates for a 6,500-megaton nuclear exchange⁶. We assumed that the zone between 30° and 70° N latitude was covered by 2×10^{14} g of smoke evenly distributed between 1 and 10 km altitude. All the complex particle injection and removal processes which would occur in the early days after the fires started are supposed to be accounted for in this value of zonal smoke loading. However, it is not well understood how particle removal processes in the initially very dense smoke plumes would affect the eventual concentration of widely distributed smoke. Early removal of smoke particles, especially through coagulation, will be more effective at higher particle concentrations. Therefore, the rates of initial horizontal dispersal and fire burning may have important roles in determining the total smoke loading a few days after the fires.

The absorption coefficient of the smoke was assumed to be $1.8 \text{ m}^2 \text{ g}^{-1}$, which gives a total absorption optical depth of 3.0. Specifically, a column density of 1.67 g m^{-2} was imposed for latitudes 31°–67° N with half this amount in the immediately adjacent zones, 27°–31° N and 67°–71° N, and no smoke elsewhere. The smoke was assumed to be evenly distributed in longitude on the basis that zonal winds would rapidly spread the smoke throughout the target latitude belt. In the vertical, equal amounts of smoke were placed in four layers of the model ranging from the 87 to 26 kPa pressure levels. Other than the overall absorption optical depth values, the details of the smoke aerosol and its optical properties are not central to our purpose of studying the first-order response of a three-dimensional circulation model to a plausible smoke aerosol. If the smoke optical depth were much greater or much less than 3, then these details of smoke optical properties would become much more important¹⁸.

The smoke appears instantaneously in all simulations at time $t = 0$ and remains fixed in place for the remainder of the model run, 20 days. This implies that a stabilized hemispheric scale aerosol is in place at day zero. (This corresponds to a time of a week or so after the fires started, essentially our $t = 0$.) Three simulations were run: 'summer' ($t = 0$ at 30 June), 'winter' ($t = 0$ at 27 December) and 'spring' ($t = 0$ at 22 March). Initial conditions for the model were supplied from a long-term (~20 yr) annual cycle simulation by R. M. Chervin.

Summertime perturbations will probably yield the largest response as more solar radiation is available for absorption by smoke. Figures 1–3 refer to the 'summer' case. The initial effect of the smoke perturbation is to cause almost all the solar flux available in the 30°–70° N latitude zone to be absorbed in the upper troposphere rather than at the surface. In the summer case, heating rates reach up to 20°C per day in the upper troposphere as soon as the smoke is injected. The absorption of solar energy by the aerosol well above the surface, combined with the escape of IR radiation from the surface layers through the smoke cloud to space, creates the surface cooling and mid-tropospheric heating which the one-dimensional models predict. (This escape of IR to space is enhanced in our three-dimensional results because nearly all upper water tropospheric clouds disappear.) Figures 1 and 2 give the zonal averages of temperature and zonal wind as a function of altitude (pressure) and latitude. These are all averaged over the period $t = 10$ –20 days and compared with an unperturbed control case.

Figure 1 shows very high atmospheric temperatures in the upper levels of the smoke and significant cooling near the surface below the smoke. The largest atmospheric warming—many tens of degrees greater than the control—is within the smoke, but

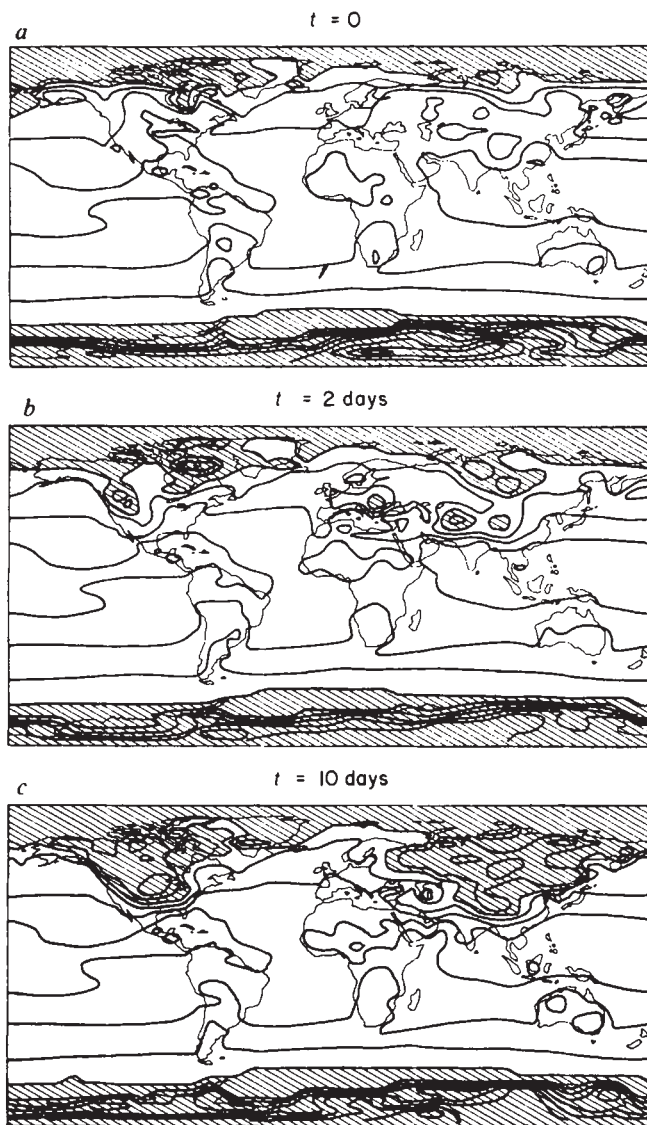


Fig. 3 Surface temperature (T) at three selected instants of time: $t = 0$ is the time at which smoke was added to the atmosphere in the summer case. Temperature contours are drawn for every 10 K. Areas with $T < 270 \text{ K}$ (that is, well below freezing) are shaded. The warmest contour value in the tropics is 300 K.

significant warming (at least 10°C) extends well beyond the fixed smoke distribution up to the 5 kPa level and southward into the Southern Hemisphere subtropics. Clearly, atmospheric motions are transporting heat beyond the area of aerosol radiative heating. Water clouds largely disappear in the middle troposphere where the heating occurs because the relative humidity decreases and little water vapour is transported upwards through the stable temperature inversion. Near the surface, cooling below the smoke occurs poleward of 30° N and exceeds 10°C (zonally averaged) between latitudes 50° and 70° N. Below 2 km altitude there is more cloud in the perturbed case.

The sharp mid-troposphere temperature gradients set up by the smoke heating support a significantly enhanced zonal circulation (Fig. 2). In the stratosphere (20–5 kPa) in the Northern Hemisphere, westerly winds are strengthened on the poleward side of the area of smoke heating and easterlies are strengthened on the equatorward side, results understandable in terms of the thermal wind relation. Substantially enhanced zonal winds exist near the surface: westerlies between 60° and 10° N, and easterlies between the Equator and 30° S. Once they have had time to develop, these enhanced winds may be expected to transport extra heat from the warmer oceans to the cooler continents.

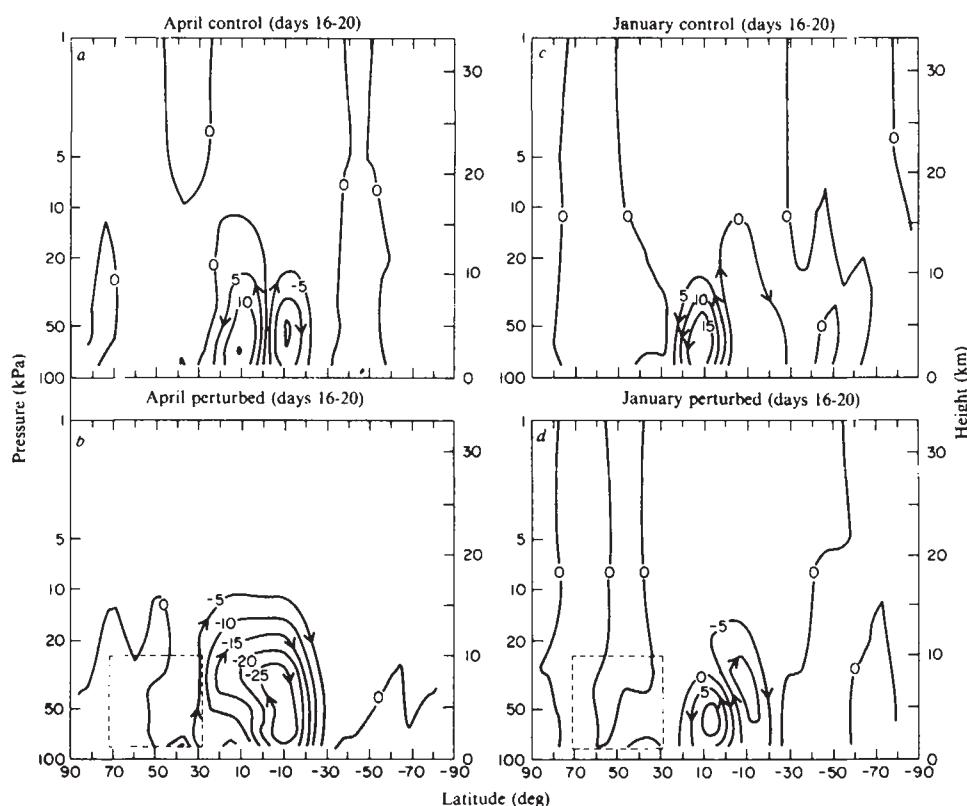


Fig. 4 Zonally averaged meridional streamfunction in units of $10^{10} \text{ kg s}^{-1}$ for the spring case (a, b) and the winter case (c, d). Arrows indicate the direction of meridional and vertical motion. Time averaging was from $t = 16$ to 20 days. The area of imposed smoke loading is indicated by the dashed box as in Figs 1, 2. The control (equilibrium simulation without smoke) and perturbed (smoke experiment) cases are shown.

Figure 3 shows the initial (30 June) surface temperatures just before smoke injection and temperatures at two subsequent times. This is, of course, a tiny fraction of the large number of 'snapshots' from the experiment which need to be examined carefully, but several important points are apparent. Even at early times ($t = 2$ days) substantial cooling has set in, with areas of subfreezing temperatures occurring beneath the smoke. This suggests that in the real atmosphere, where the aerosol would be very heterogeneous initially, rapid freezing might occur wherever dense patches of smoke drift overhead and persist for more than a few days. In that case the areas of the world that would suffer quick, transient subfreezing temperatures could be, in effect, selected at random depending on initial winds and their proximity to plumes of smoke. (Even if the smoke were evenly distributed as in our simulations, the natural variability of the weather would result in different time evolutions of regional temperatures depending on initial weather conditions at the time of the nuclear exchange.) By $t = 10$ days, temperatures would fall below freezing over substantial areas of North America and Eurasia. The coastal areas, particularly in the western parts of continents, would generally escape the intense surface cooling of continental interiors, evidently because of the enhanced advection of heat from oceans to continents discussed earlier¹⁹.

By $t = 10$ days the average land cooling under the smoke cloud for the 'summer' case is 15 to 20 °C. The ameliorating effect of the oceans is evident since temperature decreases of 30 to 40 °C would be expected on an oceanless planet⁵. Nevertheless, areas near oceans do not continuously escape subfreezing temperatures. For example, weather variability produces temperatures below freezing in Western Europe at $t = 8$ days (not shown) but not 2 days later.

The zonally averaged meridional circulation of the atmosphere is profoundly affected by the smoke perturbation. For the summer case discussed above, the normal cross-equatorial Hadley cell circulation is greatly enhanced. An even more dramatic effect is seen when the smoke injection occurs in the Northern Hemisphere spring. Figure 4a, b gives the meridional circulation for the spring case, averaged over $t = 16$ –20 days after smoke injection. The two tropical Hadley cells which normally exist (Fig. 4a) are replaced by a massive circulation

upwards and away from the smoke area. Similar results are obtained for the perturbed mean meridional circulation in the summer case. These meridional circulations may be instrumental in transporting heat away from its original site of deposition in the smoke cloud. In addition, in the real atmosphere the circulation would probably transport aerosol particles upwards and southwards, spreading them to regions beyond those in which the nuclear explosions occurred. This dynamic effect is clearly an important concern as it involves the spread of nuclear war environmental effects into areas that are not involved in the conflict^{4,5}.

Finally, in Fig. 4c, d we present the meridional circulation results for the 'winter' case, in which relatively little solar energy is available for absorption in the Northern Hemisphere mid-latitudes. Evidently there is little if any significant circulation change in the zonal average. At first sight it would appear that, in the winter case, smoke would probably not be transported out of the Northern Hemisphere mid-latitudes, but if we look at the actual three-dimensional wind patterns rather than the zonal average circulation, a different picture emerges. Streamlines of the wind at 20 kPa (about 12 km altitude) imply that smoke would be carried away from the mid-latitudes in 'streamers' for the winter (Fig. 5a) as well as for the spring (Fig. 5b) case. (Similar north-south streamlines were found in the mid-troposphere at the 50 kPa–500 mbar level.) Thus, depending on the initial north-south wind patterns, considerable amounts of smoke could be rapidly transported as far as the meteorological equator, where there would be sufficient solar radiation to create potentially large circulation perturbations. This result emphasizes the importance of using fully three-dimensional models to address the problem; the actual three-dimensional velocities associated with the circulation shown in Fig. 5b approach 50 m s^{-1} , whereas the corresponding zonally averaged velocities (see Fig. 4b) are nowhere greater than 5 m s^{-1} .

Conclusions

Our results qualitatively agree with the fundamental conclusion of the lower-dimensional models, that is, for plausible scenarios, smoke generated by a nuclear war would lead to dramatic reductions in land surface temperature. Furthermore, the three-

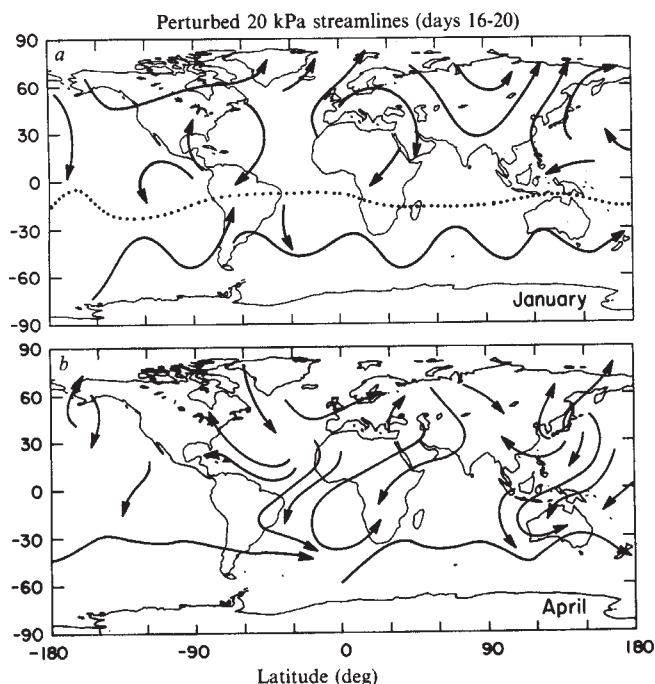


Fig. 5 Streamlines estimated from 20 kPa (200 mbar) wind vectors for spring and winter cases. The dotted line in *a* indicates the southern limit of southward winds in the winter case. If the smoke is transported to this southern limit, it would be intensely heated by sunlight and probably generate new winds which might transport it even farther.

dimensional results suggest the possibility of rapid freezing of land surfaces under transient patches of smoke that may be randomly transported by atmospheric winds. We also find significant changes in atmospheric circulation which in many cases would probably spread the smoke far beyond the altitude and latitude zones in which it was initially injected.

Clearly, further study of current model results and a greater variety of smoke injection scenarios are necessary both to analyse thoroughly physical mechanisms and to examine additional important climatic variables. Also, it should be clear that the problem is intrinsically a dynamic one. Within a few days atmospheric winds and temperature would be so profoundly altered that any estimates of aerosol spreading or removal based on today's conditions become highly questionable.

More modest improvements in model simulation should include more realistic specification of the radiative effects of aerosols, that is, inclusion of IR absorption and emission and scattering of sunlight by the aerosols. One-dimensional sensitivity studies¹⁸ indicate that inclusion of IR cooling due to smoke of visible optical depths less than ~ 10 would lead to only a small reduction in the amount of mid-atmospheric warming, and that the surface greenhouse warming would be quite small. The same studies imply that inclusion of scattering by the smoke aerosols would slightly decrease the amount of surface cooling because the aerosols will scatter some sunlight down to the surface. However, dust raised by the nuclear explosions, also not included in this study, will enhance surface cooling by backscattering sunlight so space, removing energy from the

Earth-atmosphere system^{4,5}. Moreover, such stratospheric dust or smoke scattering would also reduce the upper tropospheric heating rate for the purely absorbing smoke case, changing the calculated atmospheric circulation.

Physical processes incorporated into GCMs—including assumptions of fixed sea surface temperatures and zero land surface heat capacity, crude near-surface atmospheric representation, and sub-grid scale parameterizations for vertical and horizontal heat transport and for cloud properties—must also be critically examined. For example, vertical transport of heat by sub-grid scale processes would be affected by the dramatic increase in atmospheric stability obtained in our study. Nevertheless, our basic results for a 2×10^{14} g stabilized smoke cloud—strong land surface cooling, mid-atmospheric warming, and profound changes in circulation—seem robust; they are confirmed both by the lower-dimensional models discussed above and by results from a simplified GCM with different sub-grid scale parameterizations and with more realistic (finite) surface heat capacity²⁰. But important details such as the initial patchy freezing (Fig. 3b) are highly tentative, dependent on both the model and the initial conditions.

We believe the largest uncertainties in the nuclear aerosol/climate problem lie in translating the estimated inventory of burnable fuels in cities and forests into stabilized smoke clouds on a spatial scale suitable for global atmospheric circulation models. The way fires will burn (for example, firestorms), the height to which smoke is injected, the duration of fires, the particle concentration within the initial smoke plumes, and early particle removal by rainout in convective/mesoscale circulations all occur on spatial scales smaller than the resolution of any general circulation model now available. Unless the current estimates⁵⁻⁷ of the effect of these processes are substantially in error, however, strong cooling of mid-continental land surfaces below regional-scale smoke clouds is very plausible. Moreover, patchy, transient subfreezing outbreaks could be plausible even if hemispheric scale stabilized smoke clouds were many times smaller than the 2×10^{14} g we assumed.

Thus, the problem of long-term consequences of nuclear war represents not only an obviously critical issue for mankind, but also a stringent test of current understanding of the causes of climatic change. By subjecting models to the massive perturbation of several optical depths of aerosol, we gain insights into both model behaviour and properties of the real atmosphere which would not necessarily be as evident from studies of much smaller perturbations. Thus, we may draw implications for scientifically related problems such as the effects of volcanic eruptions on the climate and the possible massive dust injection resulting from the postulated impact of an asteroid on the Earth at the end of the Cretaceous period. It is our hope that a full hierarchy of models will be brought to bear on the question of nuclear war atmospheric effects.

We especially thank V. Ramanathan for his help in altering CCM radiation routines and for information on aerosol coagulation rates, R. M. Chervin for his annual cycle version of the CCM, J. T. Kiehl for assistance with radiation routines, and P. J. Rasch for his CCM streamfunction diagnostic package. We also thank W. Washington, M. C. MacCracken, H. N. Dalfes, P. Grutzen and A. Robock for comments on the manuscript. The National Center for Atmospheric Research is sponsored by the NSF.

Received 5 December 1983; accepted 19 January 1984.

1. Broido, A. *Bull. atom. Sci.* **16**, 409–413 (1960).
2. Ehrlich, P. R., Ehrlich, A. H. & Holdren, J. P. *Ecoscience: Population, Resources, Environment*, 690–691 (Freeman, San Francisco, 1977).
3. Crutzen, P. J. & Birks, J. W. *Ambio* **11**, 115–125 (1982).
4. Turco, R. P., Toon, O. B., Ackerman, T., Pollack, J. B. & Sagan, C. *Pap. Presented in Conf. on Long-Term Global Atmospheric and Climatic Consequences of Nuclear War*, Cambridge, Massachusetts, 22–23 April (1983).
5. Turco, R. P., Toon, O. B., Ackerman, T., Pollack, J. B. & Sagan, C. *Science* **222**, 1283–1292 (1983).
6. *Rep. of Committee on atmos. Effects of Nuclear Explosions* (National Academy of Sciences, Washington DC, in the press).
7. Crutzen, P. J. & Galbally, I. E. *Climatic Change* (submitted).
8. Raloff, J. *Sci. News* **124**, 314–317 (1983).
9. Ramanathan, V. & Coakley, J. A. *Jr Rev. Geophys. Space Phys.* **16**, 465–489 (1979).
10. MacCracken, M. C. *Pap. Presented in 3rd int. Conf. on Nuclear War*, Erice, Sicily, 19–23 August (Lawrence Livermore National Laboratory, 1983).

11. Aleksandrov, V. V. & Stenichikov, G. L. *The Proceeding on Applied Mathematics* (The Computing Centre of the USSR Academy of Sciences, 1983).
12. Washington, W. M. (ed.) *Documentation for the Community Climate Model (CCM) Version 0* (NTIS PB82-194192, National Center for Atmospheric Research, Boulder, Colorado, 1982).
13. Williamson, D. L. *Description of NCAR Community Climate Model (CCMOB)*, NCAR tech. Note TN-210+STR (1983).
14. Pitcher, E. J. *et al. J. Atmos. Sci.* **40**, 580–604 (1983).
15. Ramanathan, V., Pitcher, E. J., Malone, R. C. & Blackmon, M. L. *J. Atmos. Sci.* **40**, 605–630 (1983).
16. Sasamori, T., London, J. & Hoyt, D. V. in *Meteorology of the Southern Hemisphere*, Meteorological Monogr. 35 (American Meteorological Society, 1972).
17. Lacis, A. A. & Hansen, J. E. *J. Atmos. Sci.* **31**, 118–133 (1974).
18. Kiehl, J. T. & Ramanam, V. *Abstr. of Am. geophys. Un. Fall Meet.* San Francisco, California (1983).
19. Thompson, S. L. *et al. Ambio* (in the press).
20. Thompson, S. L. & Covey, C. *Abstr. of Am. geophys. Un. Fall Meet.* San Francisco, California (1983).

Geochronology, stratigraphy and geochemistry of Cindery Tuff in Pliocene hominid-bearing sediments of the Middle Awash, Ethiopia

C. M. Hall*, R. C. Walter†, J. A. Westgate† & D. York*

* Department of Physics, University of Toronto, Toronto, Ontario, Canada M5S 1A7

† Department of Geology, Scarborough College, University of Toronto, West Hill, Ontario, Canada M1C 1A4

Cindery Tuff is a subalkaline, rhyolitic air-fall deposit that was probably produced by a mixed-magma eruption. It is a distinctive, datable, regional isochronous marker bed within the Pliocene sediments of the Middle Awash district, and is stratigraphically situated between two new fossil hominid discoveries. Based on $^{40}\text{Ar}/^{39}\text{Ar}$ analyses of plagioclase, rhyolitic glass and basaltic glass, as well as fission-track analyses of zircons, we estimate its age to be 3.8–4.0 Myr. This implies that associated hominid skull fragments are at least 3.9 Myr old.

THE Middle Awash region contains fossiliferous terrestrial sediments that lie directly south of the Hadar area (Fig. 1). Initial geological surveys by Taieb documented abundant vertebrate fossils and thick sequences of fluvial and lacustrine sediments that range in age from Pliocene to Upper Pleistocene^{1,2}. Kalb and others undertook the first detailed geological study and recently described the stratigraphy and palaeontology of the Awash Group^{3,4}, a name which encompasses all sediments within the Awash Basin, including the Hadar Formation. Hominid fossils from the Middle Awash area were first discovered in the Middle Pleistocene Bodo Member of the Wehailtu Formation⁵. Additional hominid remains were recently found within the same beds as well as in the Lower Pliocene Sagantole Formation⁶.

We report here the results of geochronological and geochemical studies of a volcanic ash from the Lower Pliocene section of the Awash Group. Termed here the Cindery Tuff (CT), this ash is an important stratigraphic marker bed that links together isolated exposures of fossiliferous, hominid-bearing sediments along the right bank of the Awash River (Fig. 1). This is probably the same ash that Kalb and others describe as "a dark-grey, evenly-bedded, coarse to lapilli tuff" at the base of the 15 m thick Kalaloo beds^{3,7}. Clark and others recently reported the recovery of several hominid fossils from localities geographically and stratigraphically near CT⁶. A proximal femur

was recovered 7 m above CT at the Maka locality and several skull fragments were found 11 m below the tuff at the adjacent Belohdelie locality (Figs 1, 2). The Cindery Tuff not only provides excellent stratigraphic control for these hominid discoveries, but also yields feldspars, zircons and glasses that are amenable to geochronological study.

Occurrence and stratigraphy

The stratigraphic succession containing CT is one of the thickest and most laterally extensive hominid-bearing, Lower Pliocene deposits known in the world. At least 50 m of brown silts and clays, fossiliferous sands, diatomites, and occasional beds of tephra, marl and lignite lie immediately below the ash, and about 20 m of diatomites, fossiliferous sands, tephra and brown clays lie above it (Fig. 2). Sediments stratigraphically beneath CT contain a Lower Pliocene vertebrate fauna similar to the Kanapoi and Mursi faunas, while immediately above the ash, fossils contain similarities to those of the Lower Hadar Formation^{4,6}. The Cindery Tuff is at the base of the Kalaloo beds which cap the >200 m thick Sagantole Formation. The Sagantole Formation is stratigraphically between the younger Hadar and the older Adu-Asa Formations³.

We first observed the Cindery Tuff at the Bodo locality (Fig. 1). Subsequently, we found this ash at Dawaitoli, Kalaloo, Gemedah, Wilti Dora, Matabietu, Maka, Belhodelie, Wee-ee, and Sibahkaletu localities, and it can be traced in the field nearly continuously along a 25–30 km north–south exposure (Figs 1, 2).

The Cindery Tuff in this area is a 15- to 30-cm thick grey lapilli-bearing ash. It is horizontally bedded and contains several graded sequences. Angular to subrounded grey pumice clasts and black scoriaceous glass chunks (from <1.0 mm to ~5.0 mm) are supported in a fine-grained matrix (~0.125 mm) of mostly clear glass shards and dark brown glass fragments. The ash is slightly indurated and forms resistant ledges in the regional badlands topography. The lower contact is abrupt against a white diatomaceous clay; the upper contact varies from gradational to abrupt, but the superjacent lithology is everywhere a white diatomite. A conspicuous and co-extensive gastropod-bearing marl occurs 1 m below Cindery Tuff and hoof impressions were discovered on the tuff's uppermost surface at the Gamedah locality.

Analytical methods

Major element analyses were performed using a microprobe, and instrumental neutron activation analysis was used to determine rare earth element (REE) and other trace element concentrations. Samples for neutron activation were irradiated at

Table 1 Age standards

Standard	K–Ar age (Myr)	Ref.	$^{40}\text{Ar}/^{39}\text{Ar}$ age (Myr)
4B	17.29	—*	17.41 ± 0.18 17.25 ± 0.20 17.27 ± 0.30 Wtd av. 17.33 ± 0.12
4M	18.60	—*	18.74 ± 0.20
P207	83.1	35	84.61 ± 0.40 82.32 ± 0.37 Wtd av. 83.38 ± 0.27
LP6	133.6	—†	126.6 ± 4.2
	127.4	36	
	128.5	37	
MMhbl	519.5	38	529.1 ± 4.5
	518.9	37	518.9 ± 3.5 Wtd av. 522.8 ± 2.8
Fission track standard			
Fish Canyon			
Sanidine	28.5	15	
Zircon			29.07 ± 0.93§

* E. Jäger, personal communication.

† J. C. Engels, personal communication.

‡ $^{40}\text{Ar}/^{39}\text{Ar}$ ages relative to hornblende 3 gr, age 1,070 Myr.

§ Fission track age.

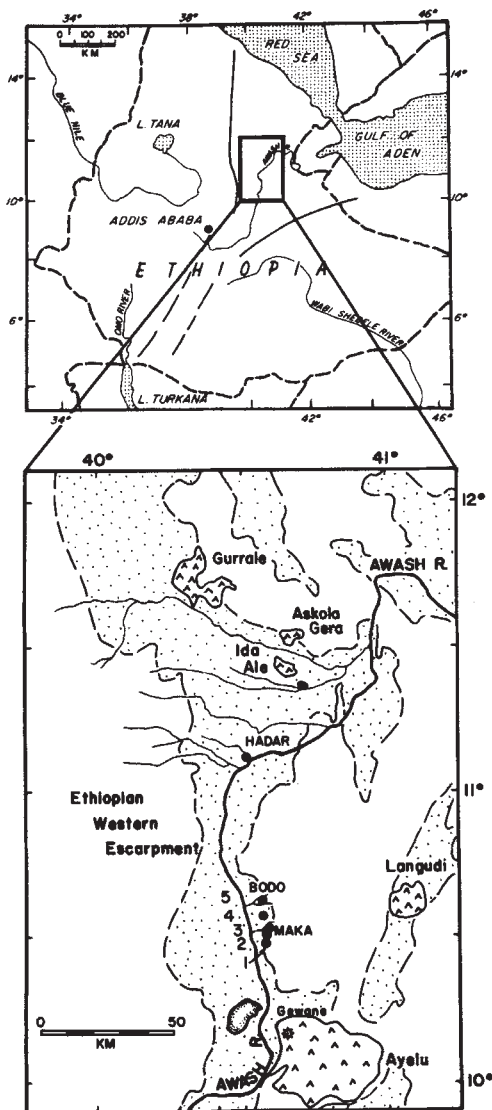


Fig. 1 Regional map of the western Afar showing the Middle Awash deposits approximately halfway between Hadar and Gewane. Filled circles (1–5) refer to stratigraphic sections in Fig. 2. Ayelu (Ayelu-Amoissu) and Langudi are Pleistocene per-alkaline rhyolite volcanoes; Ida Ale is a Pliocene subalkaline rhyolite centre and a possible source for the Cindery Tuff. Gaurale and Askola Gera are rhyolite centres of unknown age. Sediments are stippled and basalts cover the blank areas. Localities referred to in the text, but not noted above, are: Sibahkaletu, just south of locality 1; Wilti Dora, Gemedah and Kalaloo, from south to north, between localities 4 and 5; and Dawaitoli, just north of locality 5.

the University of Toronto low-flux reactor (Slowpoke II) using procedures outlined in ref. 8. $^{40}\text{Ar}/^{39}\text{Ar}$ dating methods similar to those outlined in ref. 9 were used. The primary $^{40}\text{Ar}/^{39}\text{Ar}$ standard was hornblende 3 gr (ref. 10), which was assumed to have a K–Ar age of 1,070 Myr (refs 10–13). For some shorter irradiations, our younger internal standard (age = 17.04 Myr), basalt NL1-1, which had been calibrated against 3 gr, was used. Table 1 shows the results of running a selection of international standards against 3 gr. The results demonstrate that the age we have used for 3 gr is in excellent accord with the other major inter-laboratory standards. It seems highly unlikely that there is any significant error in our $^{40}\text{Ar}/^{39}\text{Ar}$ calibrations.

Zircons from the 60–120 mesh fraction of crushed Cindery Tuff were subjected to fission track analysis using the external detector method¹⁴. The calculated neutron flux, based upon the pre-irradiated NBS SRM962 standard calibration value, was $2.59 \times 10^{15} \text{ n cm}^{-2}$, with an overall variation of less than 3%.

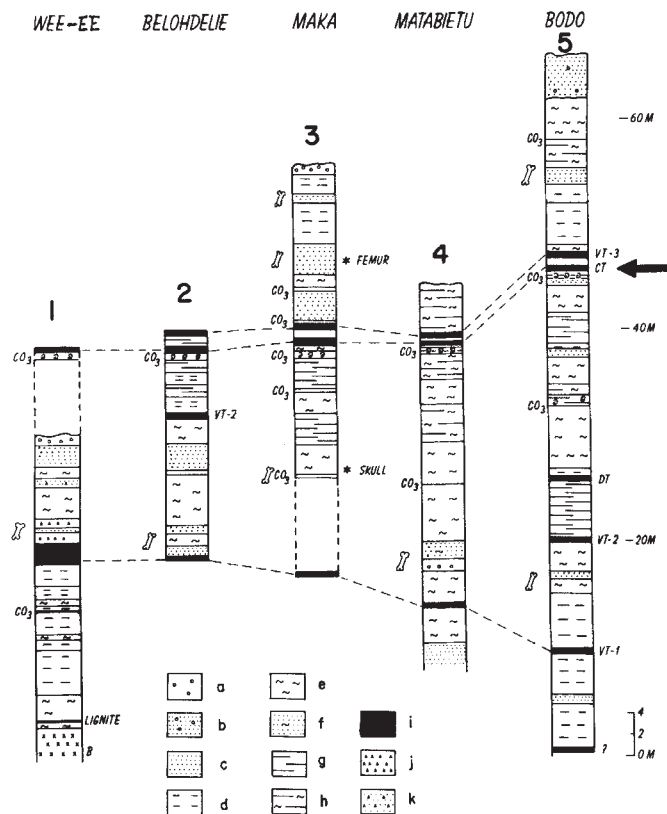


Fig. 2 Lithostratigraphic sequence of the Lower Pliocene deposits in the Middle Awash region. a, Gravel; b, gravelly sand; c, sand; d, silt; e, clay; f, sandy clay; g, diatomite; h, diatomaceous clay; i, volcanic ash; j, altered ash; k, tuffaceous sand; *, hominids; CO₂, carbonate; 'bone symbol', vertebrate fossils. Arrow designates Cindery Tuff level. Note that skull fragments were found about 700 m from the Maka femur site in the direction of the Belohdelie area. The stratigraphic context of these hominid fossils is secure with respect to Cindery Tuff, which is continuously exposed between the Maka and Belohdelie areas.

Note that the total integrated age for the Fish Canyon zircon standard in Table 1 (29.07 ± 0.93 Myr) is in very good agreement with the published sanidine K–Ar age for this sample (28.5 ± 1.3 Myr (ref. 15) calculated using the new decay constants¹⁶).

Geochemistry and petrology

Two samples of CT were selected for detailed study, one from the Bodo area (A81-1) and the other from the Maka locality (A81-37). The Bodo sample is from a locality that is well documented stratigraphically^{3–7,17}, and which serves as our reference section for the Lower Pliocene beds of the region. The Maka sample was chosen because of its proximity to two recently discovered hominid sites⁶. Therefore, geochemical parity between the Bodo and Maka samples would place the hominid and other fossil sites near the ash in a firm stratigraphic context.

A modal analysis of CT showed clear glass to be the most abundant phase, comprising 40–45%. Other phases present include 10–25% dark brown glass, 5–10% grey pumiceous chunks, 2–3% feldspar, 0.5–1% pyroxene, 0.5–1% Fe–Ti oxides and trace amounts of zircon. Basaltic and rhyolitic rock fragments comprise about 0.5% and a cryptocrystalline groundmass about 20% of the deposit.

Table 2 and Fig. 3 illustrate the close compositional similarity between the Bodo and Maka samples. The dark brown glass in both samples has a K-rich tholeiitic basaltic composition¹⁸ whereas the clear glass has a subalkaline rhyolitic composition^{19,20}. Phenocrysts of plagioclase, pyroxene and Fe–Ti oxides

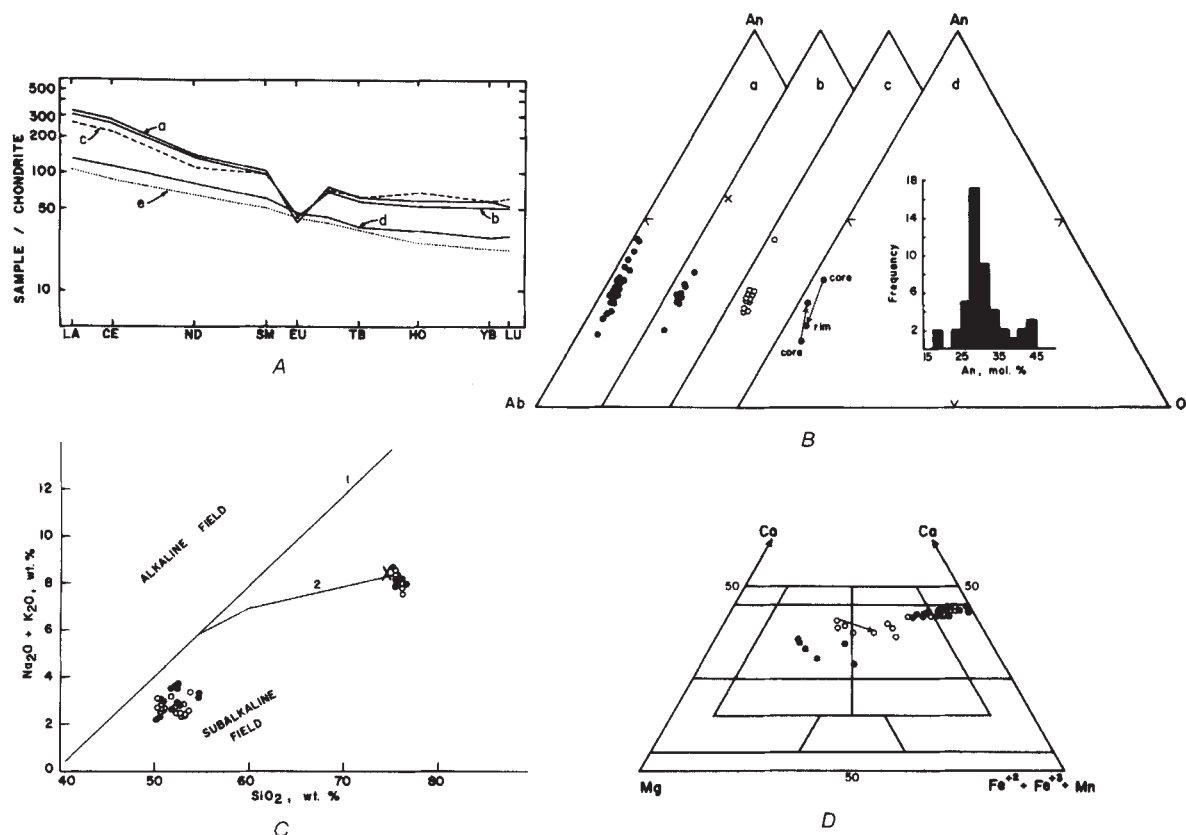


Fig. 3 A, Chondrite-normalized REE plots for Cindery Tuff rhyolitic glass (a and b) and basaltic glass (c). An Ida Ale rhyolite (dashed line, d) and an Afar series theoleiite (dotted line, e) are very similar to the CT rhyolitic and basaltic glass, respectively. Ida Ale is a likely candidate for the source of the Cindery Tuff. B, Ab-An-Or (mol%) ternary plots for Cindery Tuff feldspars: a, Maka feldspars with clear glass rims; b, Maka feldspars without glass rims; $\times$ = An55 feldspar attached to a Ti-augite grain; c, Bodo feldspars with clear glass rims; d, core-rim variations in Maka feldspars. Inset, histogram for all Cindery Tuff feldspars with attached glass, $n = 47$. C, Chemical classification of Cindery Tuff basaltic and rhyolitic glasses. Compositions were determined using a microprobe. Open circles refer to samples from Bodo (A81-1); filled circles refer to sample from Maka (A81-37). All data could not be shown because of overlap; 21 Bodo and 24 Maka rhyolitic glasses were analysed with the corresponding numbers for the basaltic glasses being 17 and 16. The $\times$ denotes the average of 28 whole rock XRF analyses for Ida Ale rhyolites²⁵. Lines 1 and 2 after refs 19, 20. D, Ternary plot for Cindery Tuff pyroxenes in the system Mg-Ca-Fe+Mn (atomic %). Symbols as in A-C.

are likewise very similar (Table 2). Chondrite-normalized REE patterns for the Bodo and Maka ash samples provide corroborative evidence for equivalence (Table 2, Fig. 3). Overall, the geochemical data strongly advocate equivalence of the sampled ash at the Maka locality to the Cindery Tuff at Bodo.

The rhyolitic glass consists of transparent vesicle-wall fragments, bubble-junction fragments, and 'stretched' pumice. They are free of microlite inclusions and there is no compositional

variation either within individual grains or between different grains or morphologies. The basaltic glass is translucent to opaque and occurs as 'fusiform', spherical scoriaceous droplets, or as angular fragments of these spheres. There is no compositional difference between these varieties and there is no compositional variation within individual grains. However, grain to grain variation is greater than that observed for the rhyolitic glass (Fig. 3). The outer edges of the translucent shards always possess thin, dark reaction or alteration rims. Phenocrysts of plagioclase, pyroxene, and Fe-Ti oxides occur as inclusions in the basaltic glass, but they are less common than acicular microlites of unknown mineralogy. Clear glass has been observed to rim several dark-brown glass spheres.

Feldspars are the most abundant mineral phase. They are generally subhedral, but their morphologies vary from euhedral to anhedral. Clear glass rims were observed on all morphologies. Zircon and Fe-Ti oxide grains are common inclusions. No optical zoning was observed in the feldspars, but microprobe traverses across several grains showed that although their interiors are quite uniform, there is occasionally an abrupt compositional shift within a hundred micrometres of the rim. The shift is significant, and can be either in a normal or reverse direction (Fig. 3).

⁴⁰Ar/³⁹Ar results

The ⁴⁰Ar/³⁹Ar age standards are listed in Table 1 and results summarized in Table 3 and Fig. 4. Three different feldspar

Table 2 Cindery Tuff fission track results

Sample	Natural tracks		Induced tracks		U (p.p.m.)	Age (Myr)
	No. tracks	Tracks per cm ² ($\times 10^6$)	No. tracks	Tracks per cm ² ($\times 10^6$)		
CF1	20	0.578	740	21.40	303	4.18
2	13	0.271	707	14.62	207	2.86
3	9	0.187	640	13.33	189	2.17
4	14	0.417	472	14.10	200	4.59
5	20	0.571	658	17.13	242	4.69
6	11	0.382	343	11.94	169	4.94
7	25	0.744	892	26.54	376	4.34
8	12	0.357	530	15.71	223	3.50
9	27	0.562	874	18.20	258	4.77
10	21	0.521	988	24.29	347	3.28
11	9	0.347	290	11.18	158	4.79
Total	181	0.438	7,130	17.27	244	3.93 $\pm$ 0.28

Neutron flux = $2.585 \times 10^{15} \text{ n cm}^{-2}$; $\lambda f = 7.03 \times 10^{-17} \text{ yr}^{-1}$; $^{235}\text{U}/^{238}\text{U} = 0.007252$; $\sigma f = 580 \text{ barn}$. Error is standard error of the mean³⁹. All analyses used the SRM 612-Cu dosimetry system.

Table 3 Cindery Tuff integrated $^{40}\text{Ar}/^{39}\text{Ar}$ data

Expt	Mass (g)	$^{40}\text{Ar}/^{36}\text{Ar}$	Vol. atmos ³⁸ ($\times 10^{-7}$)	$^{37}\text{Ar}/^{39}\text{Ar}$	J ($\times 10^{-4}$)	$^{40}\text{Ar}^*/^{39}\text{Ar}_K$ (Int.)	Int. age† (Myr)	Plateau age† (Myr)
Maka feldspars								
17	1.158	327 $\pm$ 1	7.44	3.673 $\pm$ 0.003	5.03	4.16 $\pm$ 0.12	3.78 $\pm$ 0.11	3.60 $\pm$ 0.09
18	2.312	348 $\pm$ 1	5.32	3.751 $\pm$ 0.002	5.03	4.76 $\pm$ 0.06	4.33 $\pm$ 0.06	3.91 $\pm$ 0.04
28	1.089	354 $\pm$ 2	3.63	3.761 $\pm$ 0.002	5.25	3.93 $\pm$ 0.09	3.72 $\pm$ 0.10	3.60 $\pm$ 0.13
29	1.450	322 $\pm$ 1	9.35	4.034 $\pm$ 0.002	5.25	4.19 $\pm$ 0.08	3.96 $\pm$ 0.09	3.91 $\pm$ 0.06
33‡	0.958	555 $\pm$ 9	1.27	3.070 $\pm$ 0.001	5.25	4.48 $\pm$ 0.08	4.23 $\pm$ 0.10	4.23 $\pm$ 0.09
Bodo feldspars								
48‡	4.034	581 $\pm$ 3	1.06	3.333 $\pm$ 0.001	51.1	0.440 $\pm$ 0.003	4.05 $\pm$ 0.03	3.91 $\pm$ 0.02
49‡	5.158	647 $\pm$ 4	0.80	3.374 $\pm$ 0.001	51.0	0.425 $\pm$ 0.002	3.90 $\pm$ 0.02	3.78 $\pm$ 0.02
70‡	0.753	545 $\pm$ 5	1.03	3.164 $\pm$ 0.002	47.4	0.439 $\pm$ 0.005	3.75 $\pm$ 0.04	
Bodo basaltic glass								
52	1.740	320 $\pm$ 1	11.60	2.868 $\pm$ 0.003	53.2	0.410 $\pm$ 0.005	3.93 $\pm$ 0.05	
Bodo rhyolitic glass								
53	0.955	409 $\pm$ 1	11.29	0.122 $\pm$ 0.001	52.4	0.422 $\pm$ 0.003	3.98 $\pm$ 0.03	
Feldspar age averages								
		Maka		Bodo	All			
Simple av. integrated age		4.00 $\pm$ 0.12		3.90 $\pm$ 0.09	3.97 $\pm$ 0.08			
Wtd av. integrated age		4.09 $\pm$ 0.12		3.92 $\pm$ 0.07	3.94 $\pm$ 0.05			
Simple av. plateau age		3.85 $\pm$ 0.12		3.85 $\pm$ 0.09	3.85 $\pm$ 0.08			
Wtd av. plateau age		3.90 $\pm$ 0.08		3.85 $\pm$ 0.05	3.85 $\pm$ 0.03			

† $^{40}\text{Ar}/^{39}\text{Ar}$ ages are relative to hornblende 3 gr having an age of 1,070 Myr.

‡ RF fusion. Expt 70 was a single step fusion. All errors are ± 1 s.d. Age is calculated from: $t = \ln(1 + J\{^{40}\text{Ar}^*/^{39}\text{Ar}\})/\lambda$. Constants used are those from ref. 16.

mineral separates from sample A81-37 (Maka) were analysed. The material for experiments 17, 18 and 28 was from a single separate which was the least magnetic feldspar fraction; a separate with intermediate magnetic susceptibility was fused in experiment 29, and the most magnetic fraction was analysed in experiment 33. In order to increase the age precision on subsequent runs of sample A81-1 (Bodo), a much larger mass of CT was processed for mineral separates. Experiments 48, 49 and 70 were analyses of the same feldspar separate of A81-1.

Splits of these feldspar separates used for $^{40}\text{Ar}/^{39}\text{Ar}$ geochronology were studied optically and analysed with a microprobe (Table 4). All compositions fell within the An18 to An45 range, the same as documented for those feldspars with attached clear glass (Table 4, Fig. 3).

A very pure separate of rhyolitic glass was fused in experiment 53. Optical and microprobe studies show that it contains no phenocrysts or microlite inclusions and that it yields a single compositional population (Table 4).

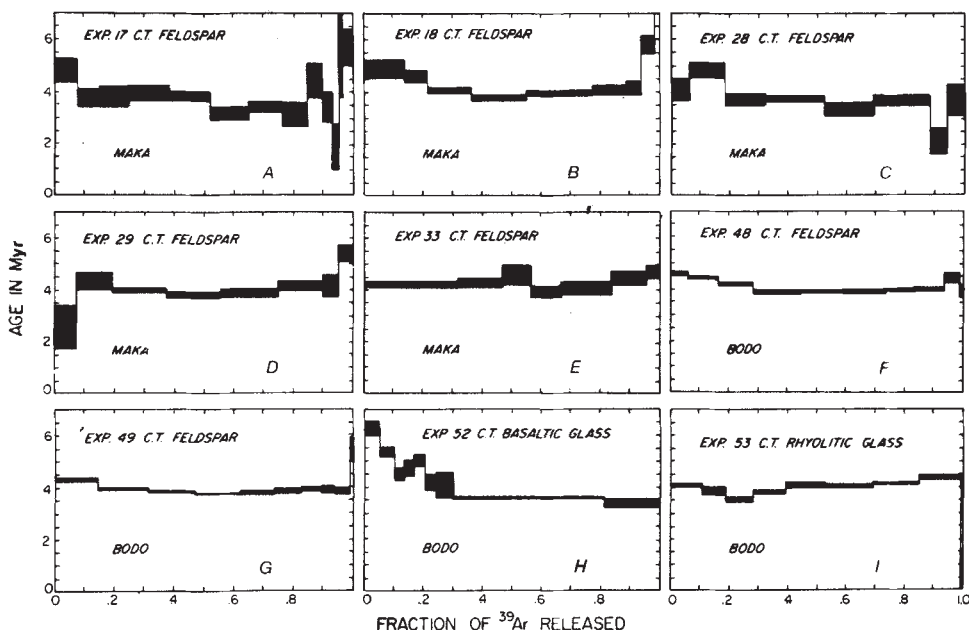
A basaltic glass concentrate was run in experiment 52. The basaltic glasses are compositionally more variable than the rhyolitic glasses and they commonly contain microlites or mineral inclusions of pyroxene, feldspar, and Fe-Ti oxides. Separates

used for dating are only about 80% pure basaltic glass due to the presence of ubiquitous included and adhering minerals. Due to the impurity of these separates, detrital or xenocrystic contamination cannot be ruled out.

The CT feldspar age spectra in Fig. 4 show a consistent pattern. There is a general drop in the apparent age in the early part of the ^{39}Ar release, with a relatively flat region in the middle of the spectrum. The large mass, higher precision runs (experiments 48 and 49) have age plateaus in the 30–90% ^{39}Ar -release region. This section of the argon release had very homogeneous $^{37}\text{Ar}/^{39}\text{Ar}$ (equivalent to Ca/K) ratios, and this was the case with all of the feldspar runs. It was decided, therefore, that the best part of the age spectra for calculating a 'plateau' age would be the 30–90% ^{39}Ar -release portion, since this represents that part of the release which is sampling the most uniform K and Ca compositions, and is thus likely to be the most representative part of the bulk mineral. The plateau ages in Table 3 are accordingly calculated from fractions which approximate, as closely as possible, to the 30–90% ^{39}Ar -release region.

No plateau age was calculated from the complex age spectrum of the basaltic glass sample (experiment 52, Fig. 4H). In contrast, the rhyolitic glass sample age spectrum (ex-

Fig. 4 Step-heating $^{40}\text{Ar}/^{39}\text{Ar}$ age spectra from the Cindery Tuff samples. A–C, Analyses of A81-37 (Maka) plagioclase separates. D, Results of an analysis of A81-37 plagioclase with clear inclusions. E, Spectrum of an A81-37 plagioclase with opaque, magnetic inclusions. F, G, Spectra from an A81-1 (Bodo) plagioclase separate. H, Analysis of an A81-1 basaltic glass concentrate. I, Spectrum from an A81-1 rhyolitic glass separate. See Table 3 for further details of these runs.



periment 53, Fig. 4I) defines a plateau over the entire ^{39}Ar release.

Table 3 gives the averages of the CT feldspar $^{40}\text{Ar}/^{39}\text{Ar}$ ages. There is no significant difference between the Maka and Bodo results, but, overall, the plateau ages tend to be slightly lower than the integrated ages. Generally, though, the feldspar ages cluster around the region of 3.8–4.0 Myr, which is in very good agreement with the rhyolitic and basaltic glass integrated ages.

Fission track results

Table 2 presents the results from the fission track analyses of zircons from the Maka site. The integrated age from the 11 CT zircon grains is 3.93 ± 0.28 Myr, which is in good agreement with the $^{40}\text{Ar}/^{39}\text{Ar}$ results. Individual grain ages were calculated to see if there was any evidence of detrital contamination in the zircon population. The standard deviation of the single grain ages, 0.93 Myr, is very close to the standard deviation expected

Table 4 Cindery Tuff (CT) glass shard and phenocryst analyses and whole rock analyses for nearby Afar volcanic rocks

	CT at Bodo locality		CT at Maka locality		Afar volcanics	
	Rhyolitic glass	Basaltic glass	Rhyolitic glass	Basaltic glass	Ida Ale rhyolites	Afar series tholeiites
$\text{SiO}_2(\%)$	75.68 ± 0.33	51.83 ± 1.67	75.82 ± 0.21	52.62 ± 1.37	74.36 ± 1.28	50.81 ± 0.66
TiO_2	0.21 ± 0.01	3.39 ± 0.23	0.22 ± 0.04	3.51 ± 0.24	0.18 ± 0.02	3.15 ± 0.29
Al_2O_3	12.30 ± 0.16	13.78 ± 0.62	12.35 ± 0.33	13.35 ± 0.19	12.80 ± 0.51	14.04 ± 0.37
FeO	2.72 ± 0.01	14.62 ± 0.74	2.66 ± 0.07	14.25 ± 0.84	2.45 ± 0.30	14.23 ± 0.30
MnO	ND	0.22 ± 0.03	ND	0.21 ± 0.04	0.06 ± 0.04	0.27 ± 0.04
MgO	ND	4.49 ± 0.75	ND	4.33 ± 0.45	0.01 ± 0.01	4.03 ± 1.18
CaO	1.00 ± 0.09	9.06 ± 0.51	1.00 ± 0.06	8.80 ± 0.60	0.84 ± 0.26	8.55 ± 0.51
Na_2O	4.14 ± 0.28	1.73 ± 0.35	4.04 ± 0.08	1.89 ± 0.48	4.85 ± 0.34	3.42 ± 0.13
K_2O	3.96 ± 0.11	0.94 ± 0.17	3.99 ± 0.04	1.02 ± 0.16	3.49 ± 0.18	0.79 ± 0.14
Cl	0.10 ± 0.01	ND	0.10 ± 0.01	ND		
P_2O_5					0.02 ± 0.01	0.58 ± 0.13
H_2O^*	6.02 ± 0.65	2.24 ± 1.17	6.44 ± 1.62	1.28 ± 0.98	$0.89 \pm$	$0.00 \pm$
(n)	(21)	(17)	(24)	(16)	(28)	(11)
Sc(p.p.m.)†	3.05	34.5	2.69		1.1	36.4
Cr	ND	12	ND		ND	15
Co	1.1	42.3	1.0		1.1	55
Zn	111	123	102		183	119
Cs	1.0	ND	1.0		1.2	ND
Ba	802	285	813		749	281
Hf	13.9	10.8	14.7		14.2	7.0
Ta	6.6	3.3	6.3		7.1	3.8
Th	17.9	5.4	18.1		16.4	4.3
U	3.9	0.9	3.7		3.4	0.6
La	101.0	40.7	95.0		82.8	33.0
Ce	221	91	211		178	70
Nd	83	§	80		73	§
Sm	20.12	11.90	18.81		19.15	9.85
Eu	2.79	3.3	3.1		3.1	3.1
Tb	3.14	1.70	2.81		3.1	1.6
Ho	4.3	2.4	3.8		5.1	1.9
Yb	11.6	5.9	10.7		12.4	4.8
Lu	1.71	0.95	1.63		1.86	0.54
(n)	(1)	(1)	(1)		(1)	(1)
	Bodo locality		Augite	Maka locality		Augite
	Ferrohedenbergite	Ferroaugite		Ferrohedenbergite	Ferroaugite	
$\text{SiO}_2(\%)$	48.24 ± 0.36	48.92 ± 0.43	50.73 ± 0.10	48.23 ± 0.73	48.86 ± 0.27	50.58 ± 0.67
TiO_2	0.40 ± 0.07	0.39 ± 0.10	0.42 ± 0.11	0.39 ± 0.08	0.45 ± 0.10	1.09 ± 0.34
Al_2O_3	0.72 ± 0.17	0.89 ± 0.28	1.25 ± 1.24	0.92 ± 0.16	1.20 ± 0.25	2.12 ± 0.54
Cr_2O_3	ND	ND	0.04 ± 0.06	ND	ND	0.15 ± 0.18
Fe_2O_3	0.57 ± 0.38	1.72 ± 1.19	1.53 ± 0.54	1.18 ± 1.34	0.74 ± 0.54	0.95 ± 0.99
FeO	28.71 ± 0.44	23.47 ± 2.55	16.47 ± 1.62	27.88 ± 1.19	25.19 ± 0.86	14.88 ± 3.90
MnO	0.75 ± 0.11	0.67 ± 0.06	0.53 ± 0.11	0.81 ± 0.13	0.81 ± 0.11	0.32 ± 0.14
MgO	1.76 ± 0.44	5.60 ± 1.63	11.23 ± 0.66	1.69 ± 0.78	4.31 ± 0.72	14.32 ± 1.77
CaO	19.57 ± 0.36	18.97 ± 0.37	17.73 ± 0.64	19.89 ± 0.26	19.62 ± 0.26	15.69 ± 1.49
Total	100.72	100.63	100.93	100.99	101.18	100.10
(n)	(10)	(5)	(3)	(10)	(4)	(6)
	Bodo locality			Maka locality		
	Oligoclase-andesine (A)	Oligoclase-andesine (B)	Oligoclase-andesine (C)	Ilmenite	Magnetite	Ilmenite
$\text{SiO}_2(\%)$	60.39 ± 1.17	60.68 ± 1.57	62.29 ± 0.78	0.39 ± 0.05	0.42 ± 0.10	0.31 ± 0.12
Al_2O_3	24.13 ± 0.76	24.45 ± 1.17	23.73 ± 0.54	49.65 ± 0.34	25.43 ± 0.48	49.97 ± 0.31
FeO	0.31 ± 0.08	0.30 ± 0.05	0.27 ± 0.04	0.32 ± 0.06	1.25 ± 0.38	0.35 ± 0.05
CaO	6.00 ± 1.07	6.48 ± 1.30	5.59 ± 1.15	4.64 ± 0.40	17.25 ± 1.20	4.40 ± 1.40
Na_2O	7.35 ± 0.48	7.09 ± 0.57	6.78 ± 0.47	44.23 ± 0.47	54.70 ± 1.28	44.15 ± 1.42
K_2O	0.56 ± 0.13	0.47 ± 0.16	0.56 ± 0.10	0.77 ± 0.06	0.62 ± 0.07	1.08 ± 0.14
Total	98.74	99.47	99.22	99.69	99.60	100.22
(n)	(12)	(24)	(7)	(5)	(5)	(5)
	Bodo locality			Maka locality		
	Ilmenite	Magnetite	Ilmenite	Ilmenite	Magnetite	Ilmenite
$\text{SiO}_2(\%)$	0.39 ± 0.05	0.42 ± 0.10	0.31 ± 0.12	0.31 ± 0.12	0.68 ± 0.26	0.31 ± 0.12
TiO_2	49.65 ± 0.34	25.43 ± 0.48	49.97 ± 0.31	49.97 ± 0.31	26.79 ± 0.36	49.97 ± 0.31
Al_2O_3	0.32 ± 0.06	1.25 ± 0.38	0.35 ± 0.05	0.35 ± 0.05	1.28 ± 0.07	0.35 ± 0.05
Fe_2O_3	4.64 ± 0.40	17.25 ± 1.20	4.40 ± 1.40	4.40 ± 1.40	11.69 ± 0.80	4.40 ± 1.40
FeO	44.23 ± 0.47	54.70 ± 1.28	44.15 ± 1.42	44.15 ± 1.42	55.30 ± 0.88	44.15 ± 1.42
MnO	0.77 ± 0.06	0.62 ± 0.07	1.08 ± 0.14	1.08 ± 0.14	0.65 ± 0.07	1.08 ± 0.14
Total	99.69	99.60	100.22	100.22	96.39	100.22
(n)	(5)	(5)	(5)	(5)	(5)	(5)

Glass and phenocryst major element analyses were made on an ETEC Autoprobe fitted with an Ortec Si(Li) detector. Microprobe operated at 20 kV with beam current set to give 3,000 counts per second on willemite; sample current about 0.23 nA; counting time 100 s. Data reduction by modified version of PESTRIPS. Total iron measured in glasses and feldspars as FeO ; Fe_2O_3 for pyroxenes and Fe-Ti oxides determined by stoichiometry. The means $\pm$ 1 s.d. are given. Whole rock analysis by XRF (see ref. 25). Feldspars A and B possess clear rims; those in group C do not, but their compositions suggest they are derived from the same population as A and B (see Fig. 3). A similarity coefficient, based on the trace element data and calculated for Bodo and Maka CT rhyolite glasses, is 0.942. ND, not detected; n, no. of analyses.

* Water by difference; glass analyses calculated on a water-free basis.

† Errors (2σ) for INAA are generally within 10%, except Ba and Cs which are within 20%.

‡ Loss on ignition.

§ Interference on Nd peak.

from track counting statistics. The fission track results agree with the $^{40}\text{Ar}/^{39}\text{Ar}$ ages in Table 3 and are consistent with the hypothesis that there is a single population of CT zircons which are about 4 Myr old.

Conclusions

The Cindery Tuff is a primary air-fall unit, which, in the Middle Awash district, accumulated in a shallow lacustrine environment. Its horizontal bedding and the manner in which it drapes the subdued, undulatory palaeotopography argue for a primary origin and the absence of obvious detrital components suggests that local reworking was not important. Its grain-size and sorting characteristics are also compatible with primary air-fall deposits²¹. Given the mixing of rhyolitic and basaltic components, we feel that CT is very likely the produce of a mixed-magma eruption²²⁻²⁴.

CT basaltic glasses are compositionally similar to Pliocene tholeiitic flood basalts from the adjacent west-central Afar volcano field²⁵ (Table 4). Furthermore, the rhyolitic glasses, even though severely hydrated, are similar to subalkaline rhyolitic lavas from Ida Ale (Fig. 3), which is less than 100 km north of the Middle Awash deposits. Preliminary grain-size analysis of CT samples from Matabietu and Sibahkaletu is compatible with a source north of Matabietu. Thus, from geochemical and geographical considerations, Ida Ale is a likely source for the Cindery Tuff.

The average of eight CT feldspar $^{40}\text{Ar}/^{39}\text{Ar}$ integrated ages is 3.97 ± 0.08 Myr ($\pm 1\sigma$) which is in excellent agreement with the integrated $^{40}\text{Ar}/^{39}\text{Ar}$ ages from both the basaltic (3.93 ± 0.05 Myr) and rhyolitic (3.98 ± 0.03 Myr) glasses, and with the zircon fission track age (3.93 ± 0.28 Myr). The feldspar plateau ages average only slightly lower at 3.85 ± 0.08 Myr, and are in general agreement with the glass and zircon results. The bulk of the geochronological evidence clearly points to an age of 3.8 to 4 Myr for the Cindery Tuff.

However, from Table 3 it is apparent that the scatter about this mean age, while not large given the low K content of the feldspar, is greater than can be accounted for by measurement error alone. It is possible that there is a variation in the argon isotope concentrations which is only averaged out when large masses are analysed. Small samples, like those used in experiments 33 and 70, are more susceptible to this variation than are the large samples, such as those in experiments 48 and 49 which had masses of 4 g or over. Detrital contamination by older material is one possible cause for such a variation. However, this seems to be an unlikely possibility. For the four different phases (plagioclase, basaltic glass, rhyolitic glass and zircon) to have the same apparent age, the tuff must have been contaminated with just the right fortuitous blend of these

materials. This consistency between phases also argues against there being any systematic excess ^{40}Ar in the feldspars. We have taken the approach that the age fluctuations in the feldspars average out when one analyses either many samples or large samples, and that the average age is a reasonable approximation to the age of the Cindery Tuff.

The low K-contents of the CT plagioclases make them less than ideal material for K-Ar dating. Inhomogeneity of K concentration within individual feldspar grains may be the cause of some of the structure in the feldspar age spectra. Internal redistribution of ^{39}Ar due to recoil effects could produce high apparent ages for the K-rich parts of the grain (low temperature) while Ca-rich regions (high temperature) would have their ages lowered. This could also be the cause of the more severe distortion in the age spectrum of the basaltic glass. The basaltic glass separate is not pure; it contains numerous small inclusions. Internal redistribution of ^{39}Ar from recoil would distort the age spectrum while leaving the integrated age unchanged²⁶⁻²⁸. It is interesting to note that the age spectra from experiments 48 and 49 have virtually the same shape as those of the KBS anorthoclase²⁹ and therefore these features may be quite common with young feldspars. The rhyolitic glass separate is clear and homogeneous, and its age spectrum is comparatively simple.

The Cindery Tuff provides a reliable age for the upper part of the Sagantole Formation^{3,7}. The Kalaloo beds, in the upper Sagantole Formation, contain fauna—particularly *Theropithecus* and *Mammuthus*—that are similar to fauna in the lower Hadar Formation⁴ but also contain some dissimilar elements³⁰. A 3.8 to 4.0 Myr age for CT is consistent with preliminary biostratigraphic correlations. Based upon these correlations, the Sidi Hakoma Member of the lower Hadar Formation is probably not older than 4.0 Myr. These data, however, only suggest that the Sidi Hakoma Member is younger than 4.0 Myr, and do not resolve the current controversy over the absolute age of the Lower Hadar³¹⁻³⁴.

Field work was done by kind permission of the Ethiopian Government, including the Ministry of Culture and Sports Affairs, the Ministry of Mines and Energy, and Addis Ababa University. Field work was supported by the US NSF through a grant to J. D. Clark, University of California at Berkeley. Laboratory work was funded through Natural Sciences and Engineering Research Council of Canada Operating Grants (J.A.W. and D.Y.) and a Collaborative Special Project Grant to J.A.W. Additional funds were supplied to D.Y. from the Provost's Fund, University of Toronto. Some travel expenses and photographic supplies for R.C.W. were kindly provided by the Canadian newspaper, *The Globe and Mail*. We thank P. Kuybida, W. J. Kenyon, D. C. Johanson, M. A. J. Williams, T. D. White, M. Tesfaye and M. P. Gorton for this invaluable help.

Received 14 July; accepted 16 December 1983.

1. Taieb, M. thesis, Univ. Paris VI (1974).
2. Taieb, M., Johanson, D. C., Coppens, Y. & Aronson, J. L. *Nature* **260**, 289-293 (1976).
3. Kalb, J. et al. *Nature* **298**, 17-25 (1982).
4. Kalb, J. et al. *Nature* **298**, 25-29 (1982).
5. Conroy, G. C., Jolly, C. J., Cramer, D. & Kalb, J. *Nature* **276**, 67-70 (1978).
6. Clark, J. D. et al. *Nature* **307**, 423-428 (1984).
7. Kalb, J. E., Oswald, E. B., Mebrate, A., Tebedge, S. & Jolly, C. J. *Newsl. Stratigr.* **11**, 95-127 (1982).
8. Barnes, S. J. & Gorton, M. P. *Geostandards Newsl.* (in the press).
9. Hall, C. M. & York, D. *Nature* **274**, 462-464 (1978).
10. Zartman, R. E. *J. Petrol.* **5**, 359-408 (1964).
11. Turner, G. *Earth planet. Sci. Lett.* **10**, 227-234 (1971).
12. Alexander, E. C. & Davis, P. K. *Geochim. cosmochim. Acta* **38**, 911-928 (1974).
13. Hanes, J. A., thesis, Univ. Toronto (1979).
14. Gleadow, A. J. W. & Lowering, J. F. *Nuclear Track Detect.* **1**, 99-106 (1977).
15. Steven, T. A., Menhert, H. H. & Obradovich, J. D. *Geol. Surv. Prof. Pap.* 575-D, 47-55 (1967).
16. Steiger, R. H. & Jäger, E. *Earth planet. Sci. Lett.* **36**, 359-362 (1977).
17. Kalb, J. et al. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **30**, 107-120 (1980).
18. Irvine, T. N. & Baragar, W. R. A. *Can. J. Earth Sci.* **8**, 523-548 (1971).
19. MacDonald, G. A. & Katsura, T. *J. Petrol.* **5**, 82-133 (1964).
20. Miyashiro, A. *Contr. Miner. Petrol.* **66**, 91-104 (1978).
21. Walker, G. P. L. *J. Geol.* **79**, 696-714 (1971).
22. Sparks, R. S. J., Sigurdsson, H. & Wilson, L. *Nature* **267**, 315-318 (1977).
23. Sigurdsson, H. & Sparks, R. S. J. *J. Petrol.* **22**, 41-84 (1981).
24. Sparks, R. S. J., Wilson, L. & Sigurdsson, H. *Phil. Trans. R. Soc.* **299**, 241-273 (1981).
25. Walter, R. C., thesis, Case Western Reserve Univ. (1980).
26. Bottomley, R. J. & York, D. *Earth planet. Sci. Lett.* **31**, 75-84 (1976).
27. Turner, G. & Cadogan, P. H. *Proc. 5th Lunar Sci. Conf.*, 1601-1615 (1974).
28. Huneke, J. C. & Smith, S. P. *Proc. 7th Lunar Sci. Conf.*, 1987-2008 (1976).
29. McDougall, I. *Nature* **294**, 120-124 (1981).
30. Kalb, J. E. et al. *J. Vert. Paleont.* **2**, 237-258 (1982).
31. Walter, R. C. & Aronson, J. L. *Nature* **296**, 122-127 (1982).
32. Brown, F. *Nature* **300**, 631-633 (1982).
33. Aronson, J. L., Walter, R. C. & Taieb, M. *Nature* **306**, 209-210 (1983).
34. Borchardt, G. A., Aruscavage, P. J. & Millard, H. T. *J. sedim. Petrol.* **42**, 301-306 (1972).
35. Lanphere, M. A. & Dalrymple, G. B. *J. geophys. Res.* **70**, 3497-3503 (1965).
36. Baksi, A. K. *Can. J. Earth Sci.* **10**, 1678-1684 (1973).
37. Roddick, J. C. *Geochim. cosmochim. Acta* **47**, 887-898 (1983).
38. Alexander, E. C., Mickelson, G. M. & Lanphere, M. A. in *Proc. 4th Int. Conf. Geochron., Cosmochron., Isotope Geol.* (ed. Zartman, R. E.) 6-8 (1978).
39. McGee, V. E. & Johnson, H. M. *Math. Geol.* **11**, 255-268 (1979).

Cryo-electron microscopy of viruses

Marc Adrian, Jacques Dubochet, Jean Lepault & Alasdair W. McDowell

European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, FRG

Thin vitrified layers of unfixed, unstained and unsupported virus suspensions can be prepared for observation by cryo-electron microscopy in easily controlled conditions. The viral particles appear free from the kind of damage caused by dehydration, freezing or adsorption to a support that is encountered in preparing biological samples for conventional electron microscopy. Cryo-electron microscopy of vitrified specimens offers possibilities for high resolution observations that compare favourably with any other electron microscopical method.

ALL biological specimens are damaged during preparation for electron microscopy. For particles in suspension, damage is caused by dehydration, adsorption onto the supporting film and by the attempts of the electron microscopist to increase the contrast. Chemical fixation, metal shadowing and negative staining are excellent methods, but they all rely on changing the specimen in order to make it more suitable for observation. As a result electron microscopy has become a highly successful science, but the material observed is generally very different from the original sample.

Cryo-electron microscopy has long been seen as a potential method for preserving the specimen in a state closer to its native state¹. The value of the approach was demonstrated by Taylor and Glaeser², who showed by electron diffraction that the periodic order of frozen-hydrated catalase crystals can be preserved down to less than 3 Å resolution, whereas it is destroyed in air-dried specimens. Another turning point was the discovery that pure water or any aqueous solution can be cooled rapidly enough to prevent formation of ice crystals^{3,4}. Based on this phenomenon, we have developed a simple method for preparing and observing frozen-hydrated specimens from native biological suspensions⁵⁻⁷.

Methodology

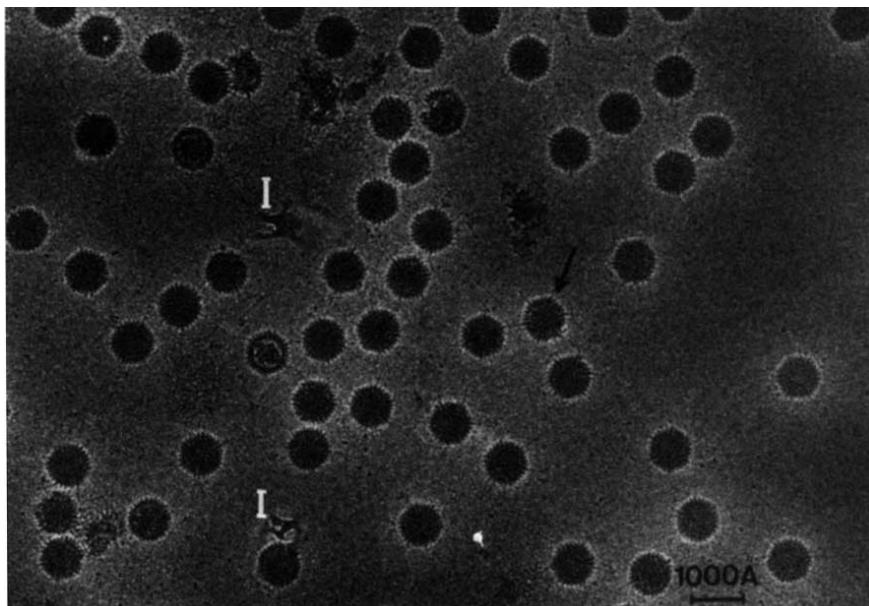
The preparation and observation of frozen-hydrated biological particles involved the following operations: (1) forming a thin layer of the suspension; (2) cooling it into the vitreous state; (3) transferring it into the microscope without rewarming above

the devitrification temperature ($T_d \approx 140$ K); and (4) observing it below T_d and with an electron dose low enough to preserve the structure of the specimen.

As discussed below, all these operations are easy to perform and take about the same time as the very simple negative staining procedure. Apart from the requirement for a cooled specimen holder, operating contamination-free below 140 K, there is no need for special instruments or installation.

The formation of a thin layer of aqueous suspension may initially seem difficult, considering the high surface tension of water, which tends to minimize the surface to volume ratio. This difficulty can be overcome, however, when the appropriate surface property is given to the supporting film^{5,8,9}. A stable thin layer forms readily in the absence of a supporting film when the liquid is stretched over the holes of a hydrophilic surface. In practice, we put a drop of solution on a clean uncoated 200-mesh copper specimen-supporting grid, remove most of the liquid with blotting paper and then freeze. On most specimen grids at least some squares are filled with a film of adequate thickness ($<3,000$ Å). Better results are obtained if the layer is allowed to become thinner by evaporation while being observed under a low-power microscope. The specimen is frozen after a few seconds, when many grid squares have the right thickness. We refer to this method as the 'bare grid method'. Figure 1 shows a specimen of adenovirus prepared in this way. An alternative way of preparation, referred to as the 'perforated film method', consists of mounting the specimen across the holes of a hydrophilic carbon film. As with the bare

Fig. 1 Thin layer of vitrified suspension of unstained and unfixed adenoviruses type 2, prepared by the bare grid method. A 5- μ l drop of suspension containing about 5×10^{11} particles per ml in hypotonic solution was put on a clean uncoated 200-mesh copper grid (Science Services, Munich) pretreated by glow discharge in air⁹. The grid was held by a tweezer mounted on a guillotine and was observed with a stereoscopic microscope at $\times 20$ magnification. Most of the liquid was removed by touching the grid edge with a blotting paper. In the next few seconds, evaporation caused the liquid layer, spread over the grid holes, to break. When about half of the grid holes were still filled, the grid was allowed to fall free from a height of 6 cm into a 1.5-cm deep liquid ethane container, cooled close to solidification temperature by liquid nitrogen. The grid was transferred to liquid nitrogen, mounted in a cold stage PW6591/100 (Philips) and rapidly introduced into a Philips 400 electron microscope where it was kept below 110 K during observation. A dozen grid squares were filled with a layer of adequate thickness. Electron diffraction confirmed that the specimen was vitrified. The micrograph was recorded with 80 kV electrons at $\times 12,500$ magnification, with a total dose of $10 \text{ e}^- \text{Å}^{-2}$ and $\sim 8 \mu\text{m}$ underfocusing. The layer is $\sim 1,200$ Å thick. Note the uniform distribution of the viruses and their well defined shape. The superposition of the upper and lower side of the viruses makes the surface structure difficult to interpret. Spikes are visible in favourable cases (arrow). Ice contaminants on the surface are marked (I).



grid method, a drop of solution is put on the grid and the specimen is frozen immediately after removing most of the liquid with blotting paper. The stabilizing effect of the perforated support ensures that a majority of holes become covered with an adequate layer of suspension. T4 bacteriophages prepared by the perforated film method are shown in Fig. 2.

Vitrification of pure water or dilute aqueous solution is achieved by rapidly cooling the liquid. With layers of less than 1 μm , it is sufficient to immerse the specimen in an efficient cryogen (for example, liquid ethane or propane but not liquid nitrogen⁵). The grid, kept under liquid nitrogen, is then transferred into the cold stage of the microscope which is in turn introduced into the electron microscope. The fact that the specimen is still vitrified in the microscope demonstrates that freezing and transfer have been done correctly.

The nature of amorphous solid water ($\text{H}_2\text{O}_{\text{as}}$) or vitreous ice remains unknown¹⁰. The fact that it can be obtained by rapid cooling of the liquid supports the idea that a continuity of state exists between the liquid and $\text{H}_2\text{O}_{\text{as}}$. The latter should thus be considered as the extreme case of supercooled water. $\text{H}_2\text{O}_{\text{as}}$ is in a metastable state, and devitrifies into cubic ice when it receives enough energy. This is the case when the sample is warmed above the devitrification temperature T_d . At a lower temperature, crystallization may take place under strong electron irradiation⁵, or when mechanical stress is applied¹¹. In the electron microscope, the layer of $\text{H}_2\text{O}_{\text{as}}$ appears as a high viscosity liquid. Particles in a broken vitrified layer can be observed flowing towards an electrically charged corner, where all the material evaporates. Heavily irradiated hydrophobic polystyrene spheres can also be seen to jump out of the film, leaving a hole which soon seals. The total absence of structure presented by a pure $\text{H}_2\text{O}_{\text{as}}$ layer at the resolution considered here is another consequence of its vitreous state. Even the focus-dependent phase contrast graininess produced by any scattering solid cannot be detected. In fact, unless some particles are added to the $\text{H}_2\text{O}_{\text{as}}$ layer, there is no difference between its image and that of a hole, exposed to the same optical density.

The contrast of frozen-hydrated specimens is low, typically three times lower than for freeze-dried specimens. If amplitude contrast only is considered, a 60 $\text{\AA}$ diameter globular protein is theoretically the smallest that can be detected with an electron dose compatible with beam damage¹². In practice, much smaller structures can be detected, as for example the fibres of the bacteriophage T4 or even DNA molecules (Fig. 2). This is because phase contrast also participates in the image formation and, in fixed beam electron microscopy, is more efficient than amplitude contrast^{13,14}. Unfortunately, the phase contrast information is not transferred equally well for all dimensions. For the large defocusing value ΔF , used in this work, optimal transfer is obtained for dimensions d such as $d^2 = 2\lambda\Delta F/(2n-1)$, whereas no transfer takes place for $d^2 = \lambda\Delta F/n$. In these equations λ is the wavelength of the electron and n is any integer. For stained specimens it is usual to choose a defocusing value that gives the best phase contrast transfer around 20 $\text{\AA}$ and to rely on amplitude contrast for much larger dimensions. This method is inadequate for unstained hydrated specimens because it leaves a gap between low resolution amplitude information and high resolution phase information. To bridge this gap, additional micrographs must be recorded with higher defocusing. Such a series may lead to a subjective image of the specimen but a precise quantitative description will require a computer-aided processing of the micrographs (not made in the present work). In Figs 1, 3, 4 and 6a, for example, the strong defocusing reveals the overall aspect of the particles, whereas fine details are revealed in the more sharply focused images of Figs 5 and 6b. The fine details seen in the strongly defocused images must be interpreted with great caution.

Electron beam damage is reduced around liquid nitrogen temperature. The cryoprotection factor, measured by the fading of diffraction points in organic crystals, is about 3 (ref. 7). The maximum dose, D_c , which can be applied without severe damage, depends on the dimension of the structure. For vitrified

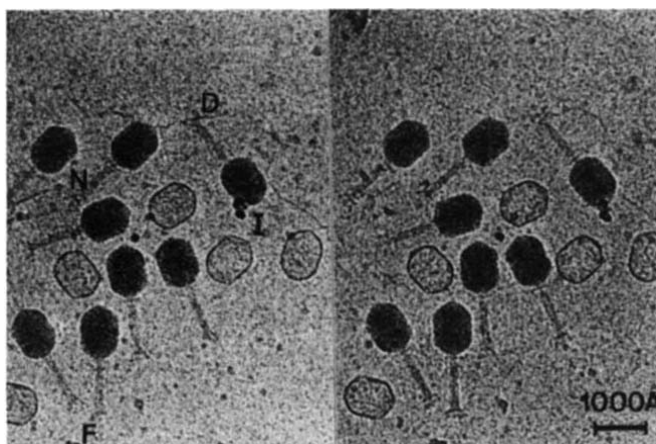


Fig. 2 Stereo view of T4 bacteriophages prepared by the perforated film method. A 5- μl drop of suspension containing about 2×10^{12} particles per ml in water was put on a perforated carbon film³¹ (average diameter of the hole 1.3 μm) mounted on a 200-mesh copper grid and treated by glow discharge in air. The specimen was prepared as for Fig. 1 except that most of the liquid was removed by firmly pressing blotting paper (Whatman; type 1) against the grid for 1 s. The field shown is from a hole of the supporting film. The layer is 1,600 $\text{\AA}$ thick. Electron optical magnification $\times 12,750$. Tilt angle $\pm 12^\circ$. The figure should be observed through a $\times 2$ magnification stereo viewer. The lower surface of the layer is revealed by contaminating ice crystals (I). The clean upper surface is not discernible. The complete bacteriophages are oriented roughly parallel to the surfaces. Shorter particles, like capsids, tails and needles (N), are more randomly oriented. Tail fibres (F) and DNA (D) are also visible.

catalase crystals it is 40 and 70 $\text{e} \text{\AA}^{-2}$ for the reflections corresponding to about 20 and 70 $\text{\AA}$, respectively. The phenomenon of bubbling, caused by the hindered escape of volatile molecular fragments induced by the beam⁵, becomes apparent at several tens of electrons per $\text{\AA}^2$. Mass-loss or the effects caused by the flow of water requires an even higher dose. Altogether, this means that high resolution images or even short tilt series can be recorded in good conditions if one of the methods usually used for low dose recording is applied.

The original distribution of the particles is generally well preserved when the unsupported vitrified layer of suspension is much thicker than the particles. The latter are then randomly positioned within the layer. When the thickness of the layer is close to the dimension of the particles, the interaction with the surface becomes noticeable. Hydrophobic polystyrene spheres tend to come out onto the surface whereas hydrophilic viruses avoid the surface by leaving the thinnest regions (adenovirus) or by orienting their long axis in a plane parallel to the surface (vesicular stomatitis virus, T4). An example is shown in the stereo view of Fig. 2.

In principle, an aqueous suspension with any concentration of solute can be used to prepare frozen-hydrated specimens. A high concentration of sugar or glycerol, or a viscous suspension, does not disturb the preparation, but a high salt concentration ($>1 \text{ M}$) leads, in some cases, to a grainy appearance. This may be due to a rearrangement of the material during freezing or to the effect of low dose irradiation. By changing the density of the solution, solutes also change the contrast of the particles. Furthermore, because some drying takes place before freezing, especially when the bare grid method is used, the final concentration of solute in the vitrified state may be significantly higher than in the original suspension. Note also that, because the particles cannot adsorb to a supporting film, and because the total volume of observed material is small, the original concentration of the particles must be high.

The excellent preservation of the particle shape is obvious in Figs 1 and 2. Figure 3 shows a vitrified preparation of the very labile preheads of bacteriophage T4, unfixed in the high-salt

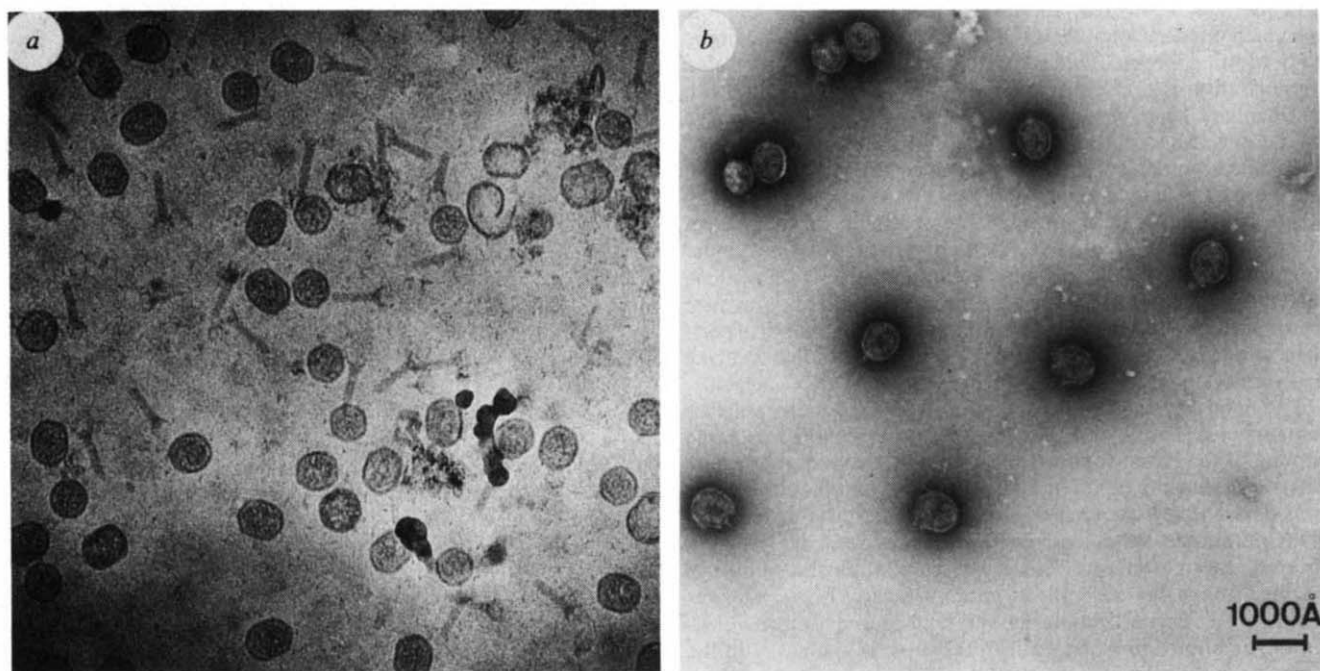


Fig. 3 Preheads of T4 bacteriophage. *a*, The prehead formed by the mutant 21⁻ and extracted by chloroform lysis in 0.5 M NaCl and one run of low speed centrifugation was prepared by the perforated film method. Electron optical magnification $\times 25,000$. The preparation shown in *b* (courtesy of B. Keller) was obtained from 21 ts mutants, fixed in 1% glutaraldehyde and further purified on a glycerol gradient. It was prepared for electron microscopy by negative staining with 2% sodium phosphotungstate. Electron optical magnification $\times 20,000$.

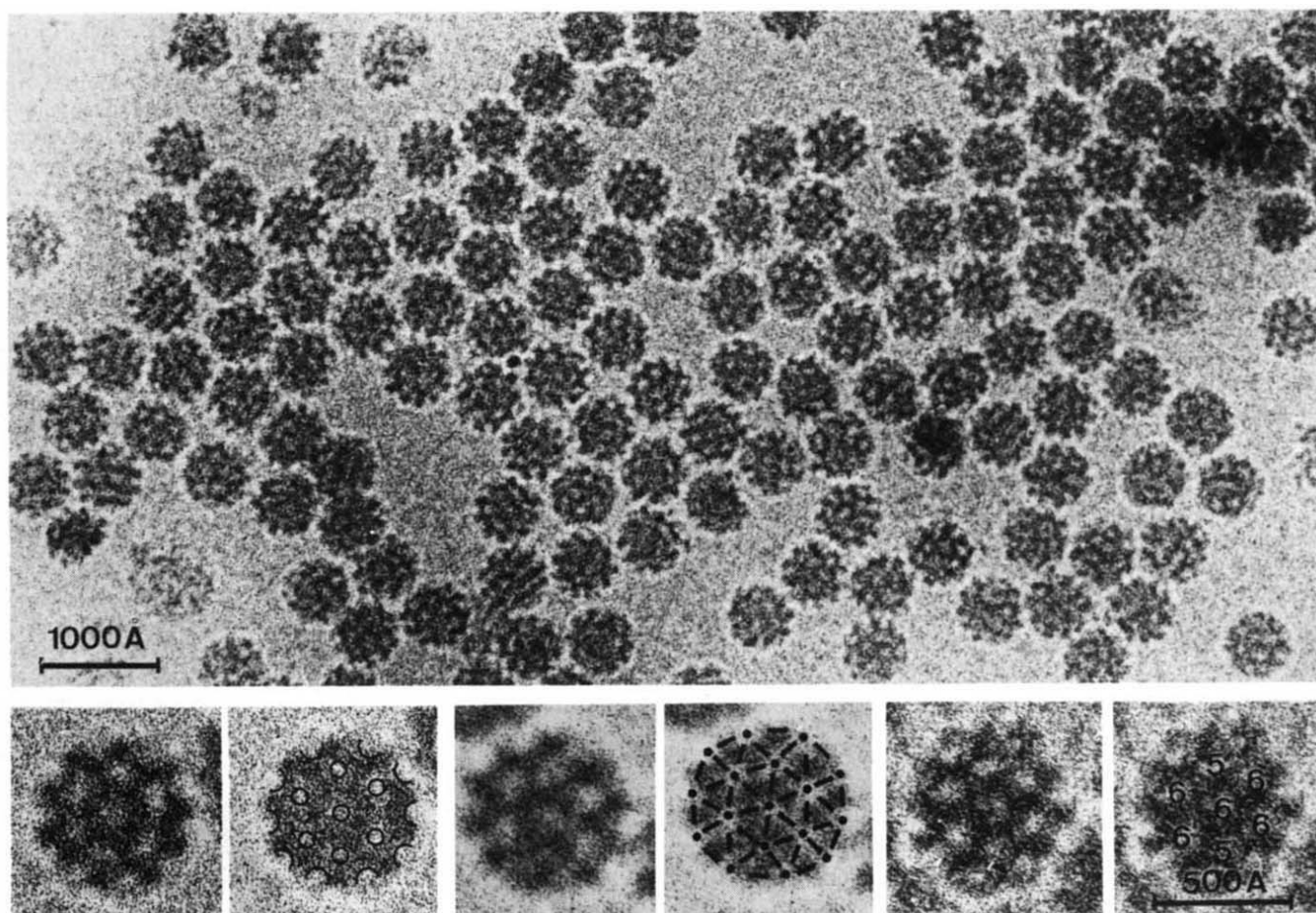


Fig. 4 Semliki Forest viruses prepared by the perforated film method. A suspension containing ~ 5 mg protein per ml in 50 mM Tris, 100 mM NaCl³², was prepared as for Fig. 2. Electron optical magnification $\times 15,000$. Underfocusing, $3.5 \mu\text{m}$. Three selected images of virions projected along their 2-fold axis are shown at higher magnification. They are underfocused by $6\text{--}9 \mu\text{m}$. Each image is duplicated and superimposed with the schematic outline of the structure, the triangular network, or a designation of the 5- and 6-fold low density nodes.

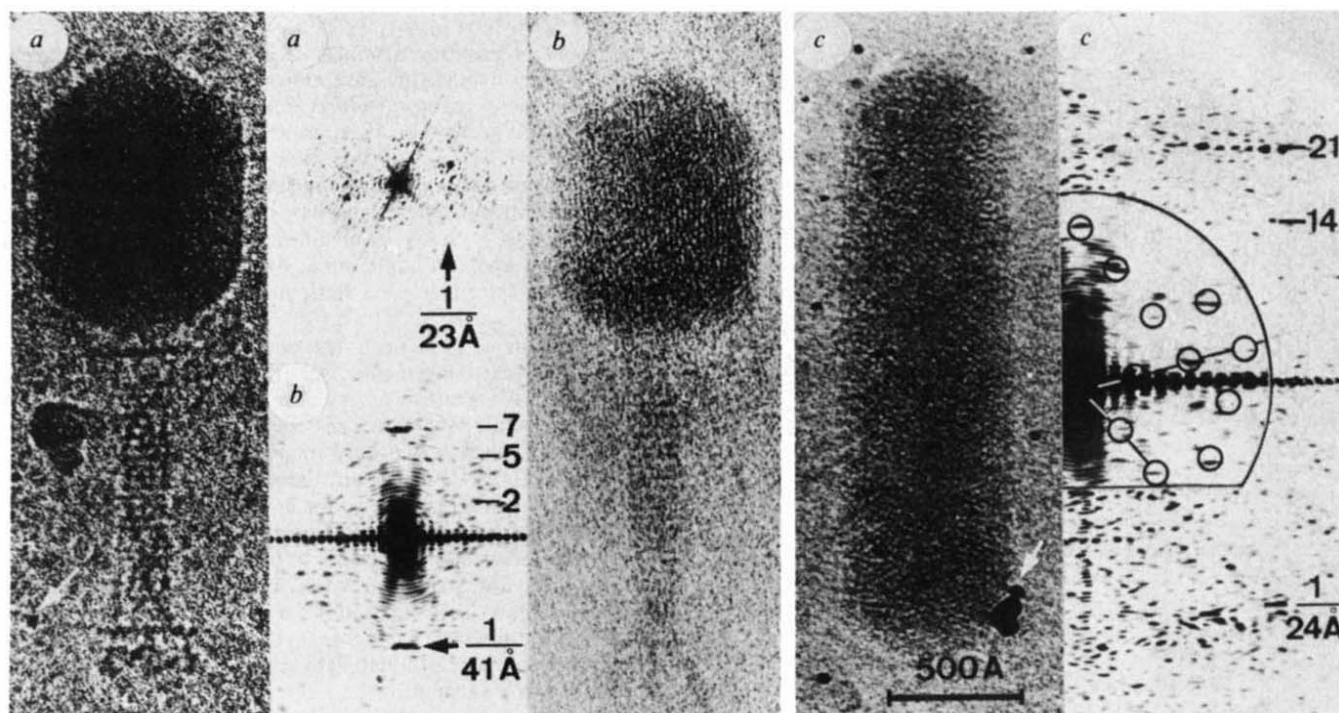


Fig. 5 T4 and CbK bacteriophages mounted on carbon films pretreated by glow discharge in pentylamine⁵. In all three micrographs the thickness of the vitrified layer is greater than that of the virus. Electron magnification $\times 42,000$. Dark spots (arrows) are gold particles deposited on the supporting film as focusing aids. The micrographs were underfocused by: *a*, 2.5 μm ; *b*, 0.5 μm ; *c*, 1.2 μm . The T4 tail structure is best seen in *a*, whereas the head structure is seen in *b*. The structure in the head of the bacteriophage CbK (*c*) is characterized in its optical diffractogram by a pair of strong reflections at $1/24 \text{ \AA}$. The central part of the diffractogram originates from a strongly underfocused micrograph (not shown). It shows the reflections due to the helical structure of the capsid. Those corresponding to one side are circled. The 14th and 21st layer lines are marked.

extraction medium (Fig. 3*a*), and fixed and desalted for the negatively stained preparation (Fig. 3*b*). Other examples of good preservation are that polyheads of the bacteriophages T4 are not flattened and the tails of the bacteriophages λ are straight.

The dimensions of the particles in the frozen-hydrated state are close to their dimensions in liquid solution. Diameters for adenovirus, λ bacteriophage and the isometric head of T4 bacteriophage are 860 ± 30 , 630 ± 20 and $850 \pm 20 \text{ \AA}$, respectively. These values agree well with the reported respective low angle X-ray diffraction measurements of 840 (ref. 15), 640 (ref. 16) and $850 \text{ \AA}$ (ref. 17). A more precise comparison can be made on organic crystals. For catalase crystals, the lattice parameters determined in water at room temperature by X-ray diffraction¹⁸ are, within 1% measurement error, identical with those measured in vitreous ice, taking ice crystals as calibration standard. This excellent agreement is surprising, as pure $\text{H}_2\text{O}_{\text{as}}$ has a 5–7% smaller density than liquid water at room temperature⁶ and would therefore be expected to expand by about 2% on vitrification. In this case also, vitrified water seems to behave like a liquid around the more rigid biological particles.

Virus structure

The structure of Semliki Forest virus remains controversial. Electron microscopical evidence obtained on negatively stained preparations suggests, although with some ambiguity, that the envelope is icosahedral with a triangulation number $T = 4$ (ref. 19). Biochemical evidence associated with independent determination of the total mass of the virion seems incompatible with this value and suggests $T = 3$ (ref. 20). As can be seen in Fig. 4, the geometry of the envelope of the Semliki Forest virus is clearly visible in frozen hydrated specimens. Because of the superposition of the two sides of the particle, only those viruses which are observed along their 2-fold axis are easy to interpret

in projection. When this is the case, two 5-fold low density nodes on the opposite sides of the 6-fold central low density node are generally visible. This configuration is the unmistakable signature of $T = 4$.

The prolate icosahedral capsid of the T4 bacteriophage is formed by two opposed caps with triangulation number $T = 13$, joined by a quasi-cylindrical segment characterized by its Q number^{21–23}. The ratio, R , of the length to the width of the capsid is directly related to Q . R varies by $\sim 3\%$ for two successive values of Q . Because of distortion during conventional specimen preparation, the value of R could not, until now, be measured accurately. When measured on micrographs similar to that of Fig. 2, R is found to be 1.431 ± 0.024 . In principle this accuracy should be sufficient to determine Q without ambiguity. We tentatively confirm the previously suggested²⁴ value of $Q = 21$ and will report elsewhere (U. Aebi, J.D., E. Kellenberger, J. Lepault and M. Wurtz, in preparation) the detailed method for measuring R and for calculating its relation to Q , taking into account that capsomeres in the caps are bent differently from those in the quasi-cylindrical segment.

High-resolution observations can also be made of frozen hydrated viruses. The structure of the tail of bacteriophage T4 is clearly revealed in Fig. 5*a*. Its diffractogram is characterized by a strong meridional reflection on the 7th layer line, corresponding to $41 \text{ \AA}$, and by strong intensities on the 2nd and 5th layer lines. This agrees with results obtained previously on negatively stained images^{25,26}. The arrangement of the DNA in the intact T4 head is more clearly visible than in previous studies. In addition to the pattern of concentric ellipses already described^{27,28}, some heads show a pattern of parallel lines separated by about $23 \text{ \AA}$ and oriented along the long axis of the head (Fig. 5*b*). In the elongated head of bacteriophage CbK, the previously undetected periodic structure of the material inside the capsid is responsible for the characteristic aspect of the image and for the strongest reflections of the optical

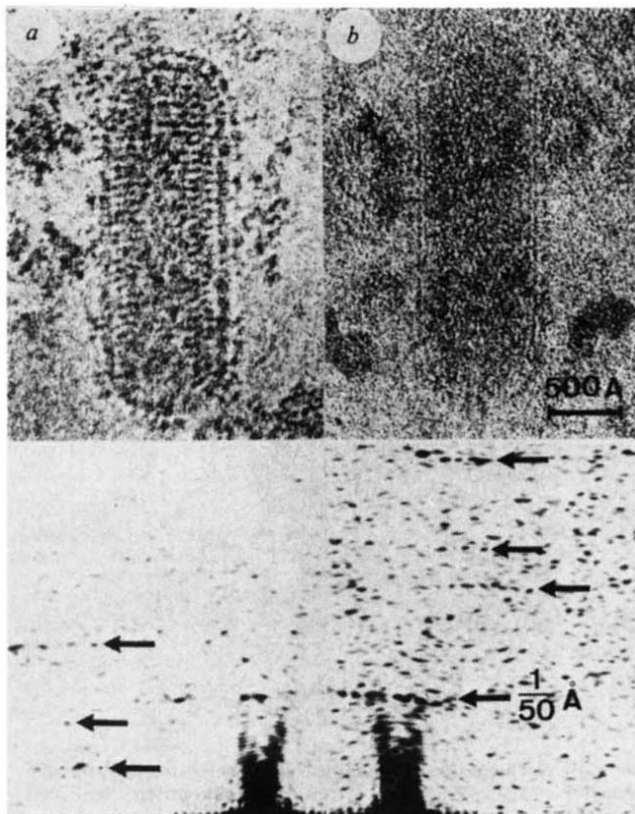


Fig. 6 Vesicular stomatitis viruses prepared by the perforated film method. Electron optical magnification $\times 43,000$. The micrographs were underfocussed by: *a*, $4\ \mu\text{m}$; *b*, $0.7\ \mu\text{m}$. The reflections which are found to be relevant, when other micrographs are also considered, are marked in the diffractograms.

diffractogram (Fig. 5c), and corresponds to a distance of $\sim 24\ \text{\AA}$ along the head axis. The lattice constant of the capsid is about $120\ \text{\AA}$ ²⁹. In frozen hydrated specimens, it is visible in strongly defocused images.

The general aspect of vesicular stomatitis virus is seen in the strongly underfocussed image of Fig. 6a, whereas the fine structure is better revealed in Fig. 6b, which is more sharply focused. The optical diffractogram reveals many more reflections than the previously observed $50\ \text{\AA}$ periodicity³⁰. Their interpretation must take into account the fact that the virus appears polymorphic, probably because most of the particles are partially degraded.

Conclusion

The usual preparation artefacts of electron microscopy seem to be largely avoided in unsupported vitreous specimens—in particular there are no artefacts due to staining, fixation and adsorption. It seems also that the thin layer of suspension is a representative sample of the bulk material, although there remains the existence of long-range interactions with the surface and a concentration effect caused by partial drying before freezing. Although at some level differences must exist between the vitreous and the liquid state, our observations show that the vitreous state is a good static model of the liquid suspension.

Fine details of biological structures are clearly revealed in frozen-hydrated specimens. All the structures previously described in the viruses we have observed have also been shown in frozen-hydrated specimens and several other important structures have been revealed here for the first time. This demonstrates once more that it is not the signal produced by a given structure which is important, but its signal-to-noise ratio. The disadvantage of the low contrast presented by frozen-hydrated specimens seems, in many cases, to be more than compensated for by their low structural noise. In particular, the reduced distortion allows us to take more advantage of periodic structures and the absence of structure in the supporting water makes the particles appear—though with low contrast—as if they were suspended in a vacuum.

Future work to establish the potential of this new electron microscopy technique should take account of the following points: (1) Water molecules seem to be mobile under the electron beam. Small molecules in solution or even biological structures could also be rearranged. This is surely the case for the molecular fragments causing bubbling. Preliminary evidence suggests that this may also be the case at low dose in concentrated salt solution. The extent to which the vitreous material can still be regarded as a static model of the liquid should be explored. (2) In some instances, the upper and lower side of a particle, in an unsupported layer, do not appear symmetrical. It is possible that the contrast may be influenced by interaction with the surfaces. (3) A strong positive staining effect occurs with biological particles when a low concentration of heavy metal salts is added to the liquid solution. Exploiting this possibility for increasing contrast could be useful in complementing the study of unstained material.

We thank Sir John Kendrew for his support during the long initial period of this research, and the following people for useful discussions and the gift of biological material: Drs U. Aebi, W. Baschong, C.-H. v. Bonsdorff, J. Caldentey, E. Kellenberger, B. Keller, K. R. Leonard, L. Philipson, V. Pirrotta, A. C. Steven, J. Tooze and M. Wurtz.

Received 8 August; accepted 21 October 1983.

1. Fernandez-Moran, H. *Expl Cell Res.* **3**, 282–350 (1952).
2. Taylor, K. A. & Glaeser, R. M. *Science* **186**, 1036–1037 (1974).
3. Brüggeller, P. & Mayer, E. *Nature* **288**, 569–571 (1980).
4. Dubochet, J. & McDowell, A. W. *J. Microsc.* **124**, RP 3–4 (1981).
5. Dubochet, J., Lepault, J., Freeman, R., Berriman, J. A. & Homo, J.-Cl. *J. Microsc.* **128**, 219–237 (1982).
6. Dubochet, J., Chang, J.-J., Freeman, R., Lepault, J. & McDowell, A. W. *Ultramicroscopy* **10**, 55–62 (1982).
7. Lepault, J., Booy, F. P. & Dubochet, J. *J. Microsc.* **129**, 89–102 (1983).
8. Hayward, S. B., Grano, D. A., Glaeser, R. M. & Fischer, K. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4320–4324 (1978).
9. Dubochet, J., Groom, M. & Müller-Neutheboom, S. in *Advances in Optical and Electron Microscopy* Vol. 8 (eds Cosslett, V. E. & Barer, R.) 107–135 (Academic, London, 1982).
10. Angell, C. A. *A. Rev. phys. Chem.* **34**, 593–630 (1983).
11. Chang, J.-J. *et al. J. Microsc.* **132**, 109–123 (1983).
12. McDowell, A. W. *et al. J. Microsc.* **131**, 1–9 (1983).
13. Erickson, H. P. & Klug, A. *Phil. Trans. R. Soc.* **B261**, 105–118 (1971).
14. Lepault, J. & Pitt, T. *EMBO J.* **3**, 101–105 (1984).

15. Devaux, C., Timmins, P. A. & Berthet-Colominas, C. *J. molec. Biol.* **167**, 119–132 (1983).
16. Earnshaw, W. C., Hendrix, R. & King, J. *J. molec. Biol.* **134**, 575–594 (1979).
17. Earnshaw, W. C., King, J. & Eiserling, F. A. *J. molec. Biol.* **122**, 247–253 (1978).
18. Unwin, P. N. T. *J. molec. Biol.* **90**, 235–242 (1975).
19. Söderlund, H., von Bonsdorff, C.-H. & Ulmanen, I. *J. gen. Virol.* **45**, 15–26 (1979).
20. Simons, K. & Warren, G. *Adv. Protein Chem.* **36**, 79–139 (1983).
21. Moody, M. F. *Virology* **26**, 567–576 (1965).
22. Aebi, U. *et al. J. supramolec. Struct.* **2**, 253–275 (1974).
23. Branton, D. & Klug, A. *J. molec. Biol.* **92**, 559–565 (1975).
24. Eiserling, F. A. in *The Bacteriophage T4* (eds Mathews, C. K., Kutter, E. M., Mosig, G. & Berget, P. B.) 11–24 (Am. Soc. Microbiol., Washington DC, 1983).
25. Amos, L. A. & Klug, A. *J. molec. Biol.* **99**, 51–75 (1975).
26. Smith, P. R., Aebi, U., Josephs, R. & Kessel, M. *J. molec. Biol.* **106**, 243–275 (1976).
27. Richards, K. E., Williams, R. C. & Calendar, R. *J. molec. Biol.* **78**, 255–259 (1973).
28. Earnshaw, W. C., King, J., Harrison, S. C. & Eiserling, F. A. *Cell* **14**, 559–568 (1978).
29. Leonard, K. R., Kleinschmidt, A. K., Agabian-Keshishian, N., Shapiro, L. & Maizel, J. V. *Jr J. molec. Biol.* **71**, 201–206 (1972).
30. Cartwright, B., Smale, C. J., Brown, F. & Hull, R. *J. Virol.* **10**, 256–260 (1972).
31. Fukami, A. & Adachi, K. *J. Electron Microsc.* **14**, 112–118 (1965).
32. Kääriäinen, L., Simons, K. & von Bonsdorff, C.-H. *Annls Med. exp. Biol. Fenn.* **47**, 235–248 (1969).

The receptor for transepithelial transport of IgA and IgM contains multiple immunoglobulin-like domains

Keith E. Mostov, Martin Friedlander & Günter Blobel

Laboratory of Cell Biology, The Rockefeller University, New York, New York 10021, USA

We have cloned and sequenced cDNA for the receptor that mediates the endocytosis and transcellular transport of IgA and IgM across many glandular epithelia into external secretions. This receptor contains five extracellular domains which are strikingly homologous to each other and to immunoglobulin variable regions.

MANY types of glandular epithelial cells synthesize a receptor that mediates transcellular transport of polymeric IgA and IgM into external secretions¹⁻³. This receptor binds these polymeric immunoglobulins (poly-Ig) at the basolateral (or in hepatocytes, sinusoidal) cell surface. The receptor-ligand complex then undergoes endocytosis and the complex is transported across the cell in vesicles and exocytosed at the apical (or in hepatocytes, bile canaliculi) cell surface⁴⁻⁶. During this process of transcytosis⁷ the ligand-binding domain of the receptor is proteolytically cleaved and subsequently discharged into the external secretion in association with the poly-Ig^{8,9}. Thus, this receptor is used for only one round of transport.

Historically, the ligand-binding domain of the receptor that is discharged with the poly-Ig was discovered first^{10,11} and termed secretory component (SC). Subsequently, several laboratories observed that SC, or a molecule immunologically related to SC, was found on the basolateral surface of epithelial cells and suggested that this molecule was the receptor for poly-Ig¹²⁻¹⁶. Using cell-free translation² and, later, pulse-chase experiments⁸ we established that SC was not synthesized as a secreted protein but as a much larger transmembrane precursor. These experiments demonstrated that SC is the proteolytically cleaved, extracellular ligand-binding domain of the poly-Ig receptor.

We report here the complete primary structure of the poly-Ig receptor deduced by cloning and sequencing DNA complementary to poly-Ig receptor mRNA isolated from rabbit liver and lactating mammary gland. The complete poly-Ig receptor primary translation product consists of 773 amino acids. Starting at the amino terminus, the structure consists of a signal peptide of 18 amino acids; an extracellular poly-Ig-binding portion of 629 amino acids; a membrane-spanning segment of 23 uncharged, predominantly hydrophobic residues and a cytoplasmic tail of 103 amino acids. The extracellular poly-Ig-binding portion consists of five highly conserved domains of approximately 100-115 amino acids. A sixth more distantly related domain follows these five and includes the membrane-spanning portion of the molecule. When the sequence of the poly-Ig receptor was compared with those of other known protein sequences, a significant degree of homology was observed between these domains and the variable (V) region of immunoglobulins as well as related proteins, such as Thy-1. Thus, the region of the poly-Ig receptor that binds immunoglobulins is itself structurally related to a portion of its ligand.

Cloning of poly-IgR cDNA

We constructed cDNA libraries from rabbit liver and lactating mammary gland, two tissues containing an abundance of poly-Ig receptor mRNA. Individual mammary clones were screened for the ability to hybridize select mRNAs coding for poly-Ig receptor polypeptides. The first clone so identified was used to screen the remainder of the mammary library by colony hybridization, and clone p32,23 was the largest selected by this procedure. We have previously observed that rabbit poly-Ig receptor has four related primary translation products: two polypeptides of

90-95,000 apparent molecular weight (MW) and two poorly resolved polypeptides of 70-71,000 apparent MW. The clone p32,23 selected mRNA coding for all four polypeptides (Fig. 1a). Northern gel analysis of both liver and mammary gland mRNA probed with nick translated p32,23 revealed hybridization with two broad bands of approximately 3.8 and 3.1 kilobases (kb) (Fig. 1b). Presumably, these represent the mRNA species coding for the four translation products.

The mammary clone p32,23 was used to identify liver clone p21,53 and both were sequenced using the Sanger dideoxy method¹⁷ (Fig. 2). p32,23 included the poly(A) tail and extended 1,563 residues from the 3' end of poly-Ig receptor mRNA. p21,53 began 362 residues from the poly(A) tail and extended an additional 3,155 residues. The two clones were identical over the 1,201 bases of overlap, in spite of having different tissues of origin. We believe that the cDNA was cloned from one of the 3.8 kb species(s) of mRNA observed on the Northern blots in Fig. 1b because the size more closely approximates that of the full length cDNA. Hence, we estimate that approximately 283 base pairs at the 5' end of the mRNA are not represented in p21,53.

Deduced amino acid sequence of poly-IgR

The complete sequence of the cloned cDNA inserts in p21,53 and p32,23 extended for 3,517 nucleotides, excluding the poly(A) tract and poly(dG) tails (Fig. 3). An open reading frame of 2,369 residues was observed to extend from nucleotides 73-2,442. The 31 amino acids coded for by nucleotides 178-271 are identical, with two exceptions, with the reported amino terminal sequence of SC determined by Edman degradation¹⁸: residue 28 is Cys, not His; and residue 31 is Pro, not Asp. The amino terminal residue of SC (Gln) is designated amino acid number 1. It is preceded by an 18 amino acid segment that begins with a Met and which is similar to the signal sequences described for many membrane and secreted proteins: 14 of the 18 amino acids are hydrophobic and the terminal residue (Ala) has a small neutral side chain¹⁹. The amino acids of the signal peptide are numbered -18 to -1. The translation initiation site is assigned to the AUG codon at the beginning of the signal peptide (nucleotides 124-126). Thus, there are 123 untranslated nucleotides on the 5' end of the open reading frame and an unusually long 3' untranslated region of 1,075 nucleotides.

There is a stretch of 23 uncharged, predominantly hydrophobic residues (positions 630-652) which are preceded by a Lys and followed immediately by the extremely basic sequence Arg-Ala-Arg-His-Arg. These features are typical of membrane-spanning segments from a variety of transmembrane proteins²⁰. On this basis, we propose that these 23 residues constitute the membrane-spanning segment of the poly-Ig receptor. This stretch also contains a single proline, indicating that the structure may kink as it crosses the membrane.

We have previously determined that the carboxy terminal end of the poly-Ig receptor constitutes the cytoplasmic tail of the protein⁸. Based on the analysis of the primary sequence data, the cytoplasmic tail of the molecule consists of 103 pre-

Fig. 1 Identification of cDNA clones for rabbit poly-Ig receptor mRNA. **a**, Translation of mRNAs selected by hybridization to p32,23 DNA. Lane 1, antiserum to rabbit SC-IgA complex (anti-SC, Cappel Laboratory) was used to precipitate the translation products of total rabbit liver RNA. Lanes 2–4, the translation products of mRNAs selected by hybridization of total rabbit liver RNA to plasmid DNAs immobilized on nitrocellulose filters. In lane 2, pBR322 DNA was immobilized on the filter and anti-SC was used to immunoprecipitate the translation products from selected mRNAs. Lanes 3 and 4 are the products obtained by immunoprecipitating the translation products of mRNAs selected by p32,23 with non-immune goat serum (lane 3) or anti-SC (lane 4). Lanes 5–7 are identical to lanes 2–4 respectively, except that total RNA from rabbit lactating mammary gland was used instead of liver mRNA. **b**, Northern blot analysis of total poly(A)-containing mRNA from rabbit liver (lane 1) or lactating mammary gland (lane 2) probed with nick translated p32,23. Positions of MW standards (*Hind*III-cut λ phage DNA) are indicated at left.

Methods: Constructing and screening of cDNA libraries. Total poly(A)-containing mRNA from rabbit liver and lactating mammary gland was prepared and size selected by sucrose gradient centrifugation². Gradient fractions enriched for mRNA coding for the 90–95,000 MW poly-Ig receptor forms, as determined by cell-free translation², were pooled. First strand synthesis was performed by incubating the following reagents in a final volume of 100 μ l at 42 °C for 1.0 h: 5.0 μ g of liver or mammary gland mRNA, 10 μ l of 10 $\times$ buffer (1.4 M KCl, 0.5 M Tris-Cl pH 8.1, 80 mM MgCl₂), 5 μ l placental RNase inhibitor (Boehringer, 30 U μ l⁻¹); 20 μ l dNTPs (2.5 mM each of all four dNTPs); 4.0 μ l oligo(dT)_{12–18} (500 μ g ml⁻¹, Collaborative Research); 4.0 μ l [³²P- α]dATP (10 mCi ml⁻¹, NEN); 5.0 μ l avian myeloblastosis virus reverse transcriptase (RVTase, 14 U μ l⁻¹, J. Beard). The reaction mixture was then boiled for 3 min, cooled on ice, centrifuged for 2 min in an Eppendorf microfuge, and the pellet discarded. Second strand synthesis using a novel procedure designed to maximize full-length synthesis was initiated by adding the following reagents: 1.0 μ l DNase-free RNase (5.0 mg ml⁻¹, Sigma); 10.0 μ l of all four dNTPs (2.5 mM); 2.0 μ l [³²P- α]dATP, and 5.0 μ l RVTase. After a 2 h incubation at 42 °C the reaction was diluted with an equal volume of H₂O. 1.0 μ l of Klenow fragment of *E. coli* DNA polymerase I (7.5 U μ l⁻¹, NEN) and 1.0 μ l RVTase were added and the mixture incubated for an additional 30 min at 37 °C. The mixture was extracted with phenol–chloroform and precipitated with ethanol. The double-stranded cDNA was cleaved with *Aspergillus oryzae* S₁ nuclease (BRL)⁵⁹. Double-stranded cDNA was size selected by sucrose gradient centrifugation and molecules greater than 1.0 kb were collected. Oligo(dC) tails were added with terminal transferase (PL Biochemicals)⁵⁹ and the poly(dC)-tailed cDNA was annealed to *Pst*I-cut, poly(dG)-tailed pBR322 plasmid DNA (NEN). CaCl₂-treated *E. coli* MC1061 (strain kindly provided by Dr K. Knight) was transformed with the recombinant plasmids⁵⁹ and approximately 600 tetracycline resistant, ampicillin sensitive colonies were obtained from mammary mRNA and 2,000 from liver mRNA. Plasmid DNA⁶⁰ from 12 of the mammary clones were individually screened by hybrid selected translation⁶¹ and one specifically selected mRNAs coding for SC precursor polypeptides. This clone contained an insert of approximately 700 nucleotides, and was used to screen the remainder of the mammary library by colony hybridization⁶². Translation and immunoprecipitation were performed as previously described². For the Northern gel analysis, 1.0 μ g of either liver or lactating mammary total poly(A) containing mRNA was separated on a 1.0% agarose gel containing 2.2 M formaldehyde⁶¹. After transfer via blotting to nitrocellulose, the samples were hybridized at 42 °C for 24 h with nick translated⁶³ p32,23 in 50% formamide, 5 $\times$ SSC (1 $\times$ SSC is 150 mM NaCl, 15 mM Na₃ Citrate) 100 μ g ml⁻¹ denatured salmon sperm DNA and 100 μ g ml⁻¹ poly(A). After hybridization the filters were washed at 60 °C with 1 $\times$ SSC followed by 0.1 $\times$ SSC and then autoradiographed with an intensifying screen.

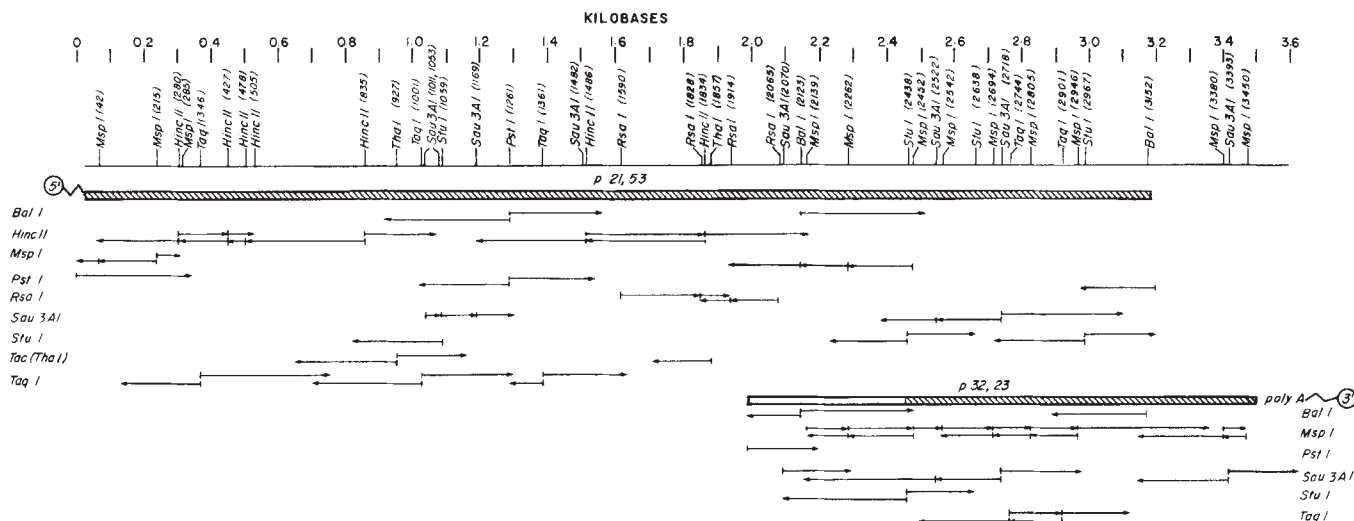


Fig. 2 Sequencing strategy. Restriction fragments of the inserts in p32,23 and p21,53 were randomly subcloned into vectors M13mp8 and M13mp9 and sequenced by the dideoxy method¹⁷. 97% of the sequence was determined in both directions and all restriction sites were sequenced across. Relevant restriction sites are indicated. Fragments are grouped according to enzyme. Arrows indicate the direction and extent of the fragment sequenced. To sequence across the *Pst*I site in the insert, a *Hinc*II clone spanning this site was isolated from a digest of intact p21,53. The M13 subclone of p32,23 used as a hybridization probe to identify p21,53 is indicated by the absence of cross-hatching. Plasmid p21,53 contained nucleotides 1–3,155, and p32,23 contained nucleotides 1,954–3,517.

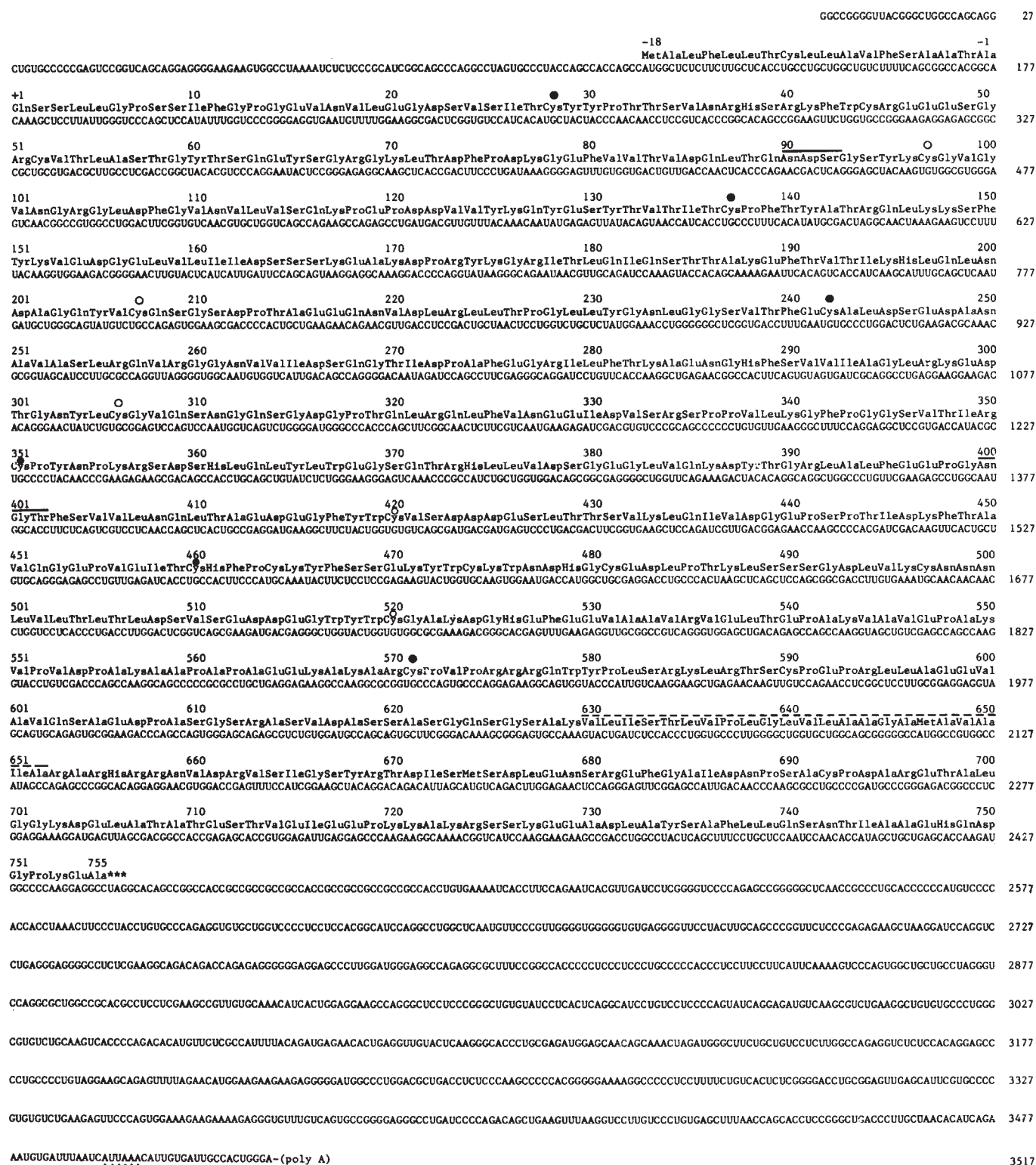


Table 1 Sequence homology between poly-IgR and other protein sequences**a, Overall sequence homology using the RELATE program**

Poly-IgR	Poly-IgR 30.6	IgK-V 11.4	Thy-1 10.7	IgH-V 7.20	IgA-C 2.46	HLA-B7 -0.14
----------	------------------	---------------	---------------	---------------	---------------	-----------------

b, Sequence homology among poly-IgR domains

Poly-IgR		Poly-IgR (s.d. unit scores)						
Domain	Segment	I	II	III	IV	V	VI	VII
I	10-118	—	7.16	6.41	7.99	8.19	0.93	0.74
II	119-223	7.16	—	6.76	6.66	2.01	2.30	-0.50
III	224-332	6.41	6.76	—	9.42	2.44	-1.72	-1.08
IV	333-441	7.99	6.66	9.42	—	6.32	1.97	-1.72
V	442-552	8.19	2.01	2.44	6.32	—	1.40	1.08
VI	553-652	0.93	2.30	-1.72	1.97	1.40	—	0.98
VII	653-775	0.74	-0.50	-1.08	-1.72	1.08	0.98	—

c, Sequence homology between poly-IgR domains and other sequences

IgK-V		IgH-V		Thy-1	
Poly-IgR segment	s.d. unit score	Poly-IgR segment	s.d. unit score	Poly-IgR segment	s.d. unit score
10-118	6.16	10-118	4.13	10-118	5.04
119-223	5.76	119-223	3.67	120-228	3.74
224-332	5.46	224-332	3.05	221-329	6.62
333-441	7.24	333-441	2.12	332-440	6.03
442-552	4.68	442-552	2.19	435-543	1.79
553-652	3.19	553-652	0.80	544-652	-1.08
653-775	-0.66	653-775	-0.92	653-755	-1.26

a, The sequence of the poly-Ig receptor (poly-IgR) was compared with itself and with IgK-V²⁴, IgH-V²⁹, rat Thy-1³⁰, IgA constant region³¹, and HLA B27³² using the RELATE program²³. The RELATE program compares all possible segments of a specified length of a given sequence with itself or another sequence. In a self-comparison, identical segments are not compared. Relatedness between segments is scored using the mutation data matrix (MDM)²⁶, which provides a probability of how often a given amino acid will either be conserved or mutated into another particular amino acid, based on studies of known protein families. Segment scores for the sequence(s) being compared are computed and compared with scores obtained by comparing 100 random permutations of the sequence(s). The final score is the difference between the score for the sequence(s) being compared and the average of the scores for the comparisons among randomized sequences, divided by the standard deviation (s.d.) score of the randomized sequences. This score, in s.d. units, is thus an overall indication of either the degree of internal homology of a sequence compared with itself, or the relatedness of two sequences being compared. **b,** The poly-Ig receptor was divided into seven segments and these were compared with each other using the ALIGN program. This program finds the best possible alignment between two given sequences of similar length, using the MDM for scoring and assigning a penalty for gaps. Much like the RELATE program, the ALIGN program compares the score of the optimal alignment of the sequences being compared with the scores of the optimal alignments of 100 random permutations of the sequence. The final score, in s.d. units, is the difference between the optimal score of the sequences being compared and the average score of the comparisons among random permutations, divided by the s.d. of the scores of the randomized sequences. **c,** The segments of poly-Ig receptor were compared with IgK-V, IgH-V, and Thy-1 using the ALIGN program. Data in this table were generated in the laboratory of J. Maizel.

dominantly hydrophilic amino acids with a total MW of 11,120. This is consistent with our previous estimates that ~15,000 daltons at the carboxy terminus of poly-Ig receptor are accessible to proteolytic digestion from the cytoplasmic face of the membrane. In collaboration with D. Osterman an antiserum has been prepared to a synthetic polypeptide comprising a 17 residue portion of the tail region of poly-Ig receptor. Using this antiserum we have found that this tail can be proteolytically cleaved from the cytoplasmic face of the membrane and recovered as a soluble, non-membrane associated fragment (in preparation).

The site at which poly-Ig receptor is cleaved to produce SC and a residual fragment containing the transmembrane segment plus the cytoplasmic tail portion is not known. However, the carboxy terminal amino acid sequence of rabbit SC is Ala-Glu as determined by carboxypeptidase Y digestion¹⁸. The Ala-Glu dipeptide occurs three times (residues 563-564, 597-598 and 605-606) in the region just upstream from the transmembrane segment, which is where we predict cleavage of the poly-Ig receptor to SC occurs. Hence, it remains to be determined which of these sites is the correct cleavage site. Interestingly, all three of these sites would require cleavage between two acidic residues. The MW for poly-Ig receptor, including the signal sequence, was calculated to be 82,041. As discussed above, there are two poly-Ig receptor translation products of 90-95,000 and two of 70-71,000 apparent MW. Considering the inaccuracies of MW estimation by SDS-polyacrylamide gels, it is most likely that we have cloned cDNA for an mRNA species coding for one of the 90-95,000 apparent MW polypeptides, since these are closer to the predicted MW.

Homologous repeating units

The extracellular portion of poly-Ig receptor consists of 629 amino acids. We noted that it contained 18 cysteines, 10 of which could be grouped into regularly spaced pairs (residues 28 and 97, 137 and 207, 242 and 306, 351 and 420, 460 and 520) in which the members of each pair are separated by 60 to 70 amino acids. The sequences surrounding the corresponding members of each pair have a high degree of homology, suggesting that the poly-Ig receptor contains 5 repeating units, each approximately 100-115 residues long.

We confirmed this by computer-assisted analysis. All of the computer studies were performed on a VAX 11/750 computer in the laboratory of Dr Jacob Maizel. The internal repeating domains of the poly-Ig receptor are dramatically illustrated in the dot matrix analysis^{21,22} shown in Fig. 4a. The solid diagonal line bisecting the figure is the result of identity of the poly-Ig receptor with itself. The broken diagonal lines running parallel to it reveal the internal repeats. Shifts in the lines are due to insertions or deletions. This analysis strongly suggests that the extracellular portion consists of five repeating units of 100-115 amino acids. There are also a number of shorter diagonal lines scattered in the figure, suggesting that there may be several smaller, less constant repeating units within these larger 100-115 amino acids repeats.

We quantitated the degree of homology between domains using the RELATE and ALIGN computer programs²³. The RELATE program gives an overall measure of either internal homology of a sequence when compared with itself, or of homology between two sequences. The final score is in standard

deviation (s.d.) units and reflects the probability that the observed degree of homology could be due to chance. A score of 3 s.d. units, for instance, indicates a degree of relatedness that has a probability of 0.135×10^{-2} of occurring by chance, and is taken to indicate a statistically significant result. The poly-Ig receptor compared with itself has a score of 30.6, indicating a high degree of overall internal homology (Table 1a).

We used the ALIGN program to determine the extent of homology between the individual domains. This program finds the best possible alignment between two given sequences of similar length. Much like the RELATE program, the final score is given in s.d. units, and reflects the probability that the observed homology could be due to chance. We divided the poly-Ig receptor segment into seven roughly equal segments. We initially chose the boundaries between the first five segments by extrapolating the homology lines of the dot matrix analysis back to the horizontal axis. The boundaries of the sixth segment were based on homology with the immunoglobulin κ chain V region (see below). As seen in Table 1b, the first five domains share significant homology with each other, although to variable degrees. By both the dot matrix and ALIGN analysis, the sixth segment (which includes the membrane-spanning region) and the seventh segment (the cytoplasmic tail) do not share significant homology with the five repeating units.

Homology with immunoglobulins

The similarity in length between the poly-Ig repeating units and those of immunoglobulins and the similar positioning of the paired cysteine residues²⁴ led us to compare the sequence of the poly-Ig receptor with a number of immunoglobulins and immunoglobulin-like proteins. We used the LSRCF program²⁵ to compare the poly-Ig receptor amino acid sequence with the entire Dayhoff Protein Data Bank²⁶ and used the related program SRCHN²⁵ to compare the poly-Ig receptor nucleotide sequence with the Genbank Data Bank²⁷. Homologies were found between poly-Ig receptor and numerous members of the immunoglobulin superfamily, with the best scores being with immunoglobulin κ chain variable regions. Homologies were most significant at the amino acid level, suggesting that evolutionary constraints were placed predominantly at the level of protein structure, rather than nucleotide sequence. Figure 4b shows the dot matrix comparisons of poly-Ig receptor with three representative high scoring proteins: immunoglobulin κ chain variable region²⁸ (Ig K-V, human SCW), immunoglobulin heavy chain V region²⁹ (Ig H-V, mouse AMPC1), and the rat Thy-1 antigen³⁰. We used the RELATE program to give an overall measure of relatedness between poly-Ig receptor and other proteins (Table 1a). Highly significant scores were obtained for Ig K-V, Thy-1 and Ig H-V. By way of comparison, the constant region of IgA³¹ (human IgA1 Bur) gave a borderline score of 2.46. HLA B7³², which is known to be related to the immunoglobulins, gave a negative score. We used the ALIGN program to compare quantitatively specific domains of the poly-Ig receptor with the various other molecules (Table 1c). Of note is that the sixth domain of the poly-Ig receptor shows significant homology with Ig K-V, both in the dot matrix analysis (Fig. 4b) and by the ALIGN score (Table 1c). The dot matrix analysis even suggests small stretches of homology between a portion of the seventh segment and Ig K-V, Ig H-V and Thy-1.

The optimal alignments generated by the ALIGN program are not presented, but indicated that the poly-Ig receptor domains have an average of 28% identical residues with Ig K-V. The per cent homology for each segment is I, 36; II, 26; III, 31; IV, 23; V, 30; VI, 23. When comparisons are based on grouping amino acids on the basis of similar side chains, the extent of homology increases to an average of 56%. The per cent homology for each segment is I, 59; II 62; III, 56; IV, 51; V, 61; VI, 49. To a large extent, the residues which are most conserved among the five poly-Ig receptor domains are also relatively invariant among the κ chain variable sequences³³. The best example of this is an 18 residue segment which has the

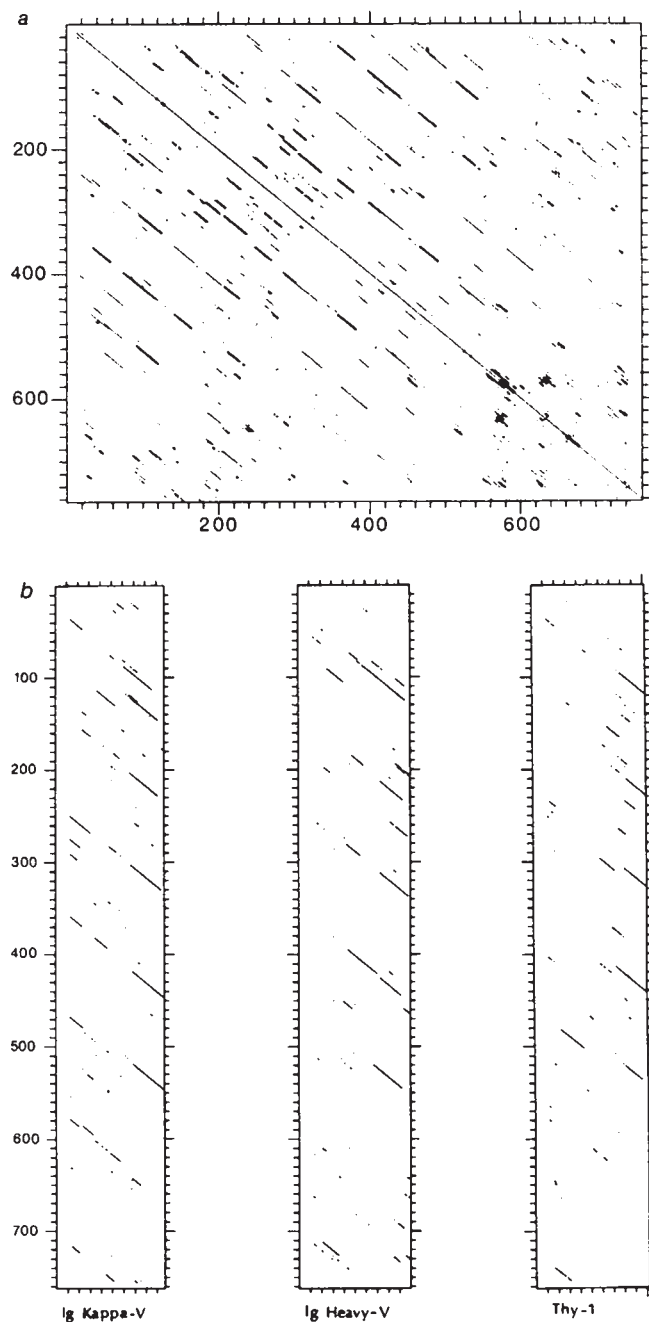


Fig. 4 Dot-matrix sequence analysis. *a*, Self-comparison of poly-Ig receptor. The complete amino acid sequence of the poly-Ig receptor is presented on both the horizontal and vertical axes. The 'DOT-MAT' program developed by David George of the National Biomedical Research Foundation compares two sequences to allow visual detection of homology. Defined length segments of residues from the sequence represented on the vertical axis are sequentially compared with segments of the sequence on the horizontal axis, and homologies between them are scored using the mutation data matrix (MDM). In this figure, segments of 25 residues from the vertical axis were compared with segments from the horizontal axis, and a dot is placed at the appropriate position when the total MDM score for the comparison is 15 or greater. By adjusting the size of the segment and the minimum score needed for a dot, the stringency of the homology comparison can be varied. (See refs 21 and 22 for a more extensive discussion of dot matrices.) The amino acid sequence of the poly-Ig receptor is numbered sequentially along the axes, with the first residue of the signal sequence as amino acid number 1. *b*, Comparison of poly-Ig receptor with immunoglobulin variable regions and Thy-1. In each of the three panels, the amino acid sequence of the poly-Ig receptor is represented on the vertical axis and the amino acid sequence of the other protein on the horizontal axis. (This figure was generated in the laboratory of Dr J. Maizel.)

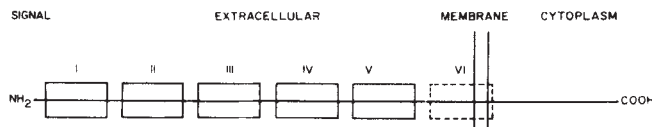


Fig. 5 Schematic representation of poly-Ig receptor domain structure. The five highly conserved domains in the extracellular portion of the molecule are indicated by solid-bordered rectangles. The sixth domain is indicated by an interrupted rectangle. The lipid bilayer of the membrane is indicated by the two vertical lines, and the locations of the segments with respect to the membrane are indicated above.

consensus sequence Phe-Ser-Val-Thr-Ile-X-X-Leu-X-X-Glu-Asp-X-Gly-X-Tyr-X-Cys and which is found in the poly-Ig receptor sequence at positions 80–97, 190–207, 289–306, 403–420 and 503–520. This segment resembles the proposed primordial building block of the immunoglobulin variable region³⁴.

Considering these sequence homologies and the conserved pairs of cysteine residues, it is very likely that the poly-Ig receptor domains share a common ancestry with the Ig domains. It is reasonable to suggest that the first five poly-Ig receptor domains have a three-dimensional structure similar to that observed for the Ig domains³⁵. We also predict that the conserved pairs of cysteines are disulphide linked. Apparently the sixth domain, while sharing sequence homology with the immunoglobulin domains, has evolved somewhat differently. Although it appears to have the first Cys (position 571) of the conserved pair of cysteines, it has lost the second member of the pair. Moreover, its carboxy terminus has apparently been replaced with a membrane spanning segment. Interestingly, the first portion of the membrane spanning segment (residues 648–655, Val-Leu-Ile-Ser-Thr-Leu-Val-Pro) is homologous to a segment of Ig K-V (residues 73–80, Leu-Thr-Ile-Ser-Thr-Leu-Gln-Pro), suggesting that part of the membrane spanning segment of the receptor may also have evolved from the same ancestor as the immunoglobulin domain.

The concept that receptors for proteins have structural similarities with their ligands has previously been proposed for cobalamin receptor and other receptors³⁶. Our sequence data substantiate this principle and extend it to the immune system.

The similarity of the poly-Ig receptor domains to immunoglobulin variable regions suggests that the poly-Ig receptor may have evolved from an immunoglobulin (or immunoglobulin-like protein) which had a variable region with specificity for a site on polymeric IgA and IgM. The poly-Ig receptor may be a prototype for other immunoglobulin receptors (for example various leukocyte Fc receptors^{37,38} or the receptors for trans epithelial transport of IgG in placenta³⁹ and neonatal intestine⁴⁰) and perhaps for other polypeptide receptors (for example the insulin receptor⁴¹), which may have evolved in a similar fashion. Several classes of molecules involved in recognition in the immune system, such as immunoglobulins, histocompatibility antigens⁴², and perhaps the T-cell receptor⁴³, all have the same basic immunoglobulin domain. Furthermore, the Thy-1 antigens, and perhaps other molecules which may be involved in cell-cell interactions and development, also have the same structure⁴⁴. Thus the immunoglobulin domain appears to be a fundamental, yet versatile, structure utilized in a wide variety of receptor-mediated recognition processes.

An alternative explanation of the observed homology between the poly-Ig receptor and immunoglobulin domains is that the interaction between the poly-Ig receptor domains and the domains of the immunoglobulin ligand somehow resembles the interactions of the immunoglobulin domains with each other. IgA and IgM form polymers (even in the absence of the J chain⁴⁵) so their domains must have the ability to bind to each other.

In addition to a noncovalent association, the secretory component forms two disulphide bonds with some, but not all, subclasses of IgA^{1,46}. In human SC, the sequences surrounding the two cysteines involved are His-Phe-Cys-Lys-Pro-Gln-Tyr and Phe-Val-Asn-Cys-Glu-Asp⁴⁷. The cysteines involved in

rabbit SC are not known, but in the rabbit sequence the closest homologues are the sequences surrounding the cysteines at position 464 (His-Phe-Pro-Cys-Lys-Tyr-Phe) and position 481 (Asn-Asp-His-Gly-Cys-Glu-Asp) respectively. These homologies suggest that it may be the fifth domain of the poly-Ig receptor that becomes disulphide linked to its ligand. The locations of the cysteines in IgA which are involved are unknown. Although SC becomes disulphide linked to only one of the monomers of dimeric IgA⁴⁸ it is possible that the other SC domains interact noncovalently with both IgA monomers and may stabilize the complex^{49,50}. Disulphide bonds are apparently also formed between insulin and its receptor⁵¹.

Rabbit SC consists of four primary translation products while human SC consists of only one. At least part of the rabbit heterogeneity is genetic, since two alleles of rabbit SC are known⁵² and most rabbits, including the ones used in this study, are heterozygotes. However, even homozygous rabbits have two forms of SC, differing in apparent MW by about 20,000¹⁹. The amino terminal 31 amino acids and the carboxy terminal two amino acids of these two SC polypeptides have been shown to be identical, suggesting that the smaller polypeptide arose by an internal deletion of the larger¹⁸. There could be two genes coding for the two polypeptides. Alternatively, by analogy with immunoglobulin heavy chains⁵³, the transcript of one gene could be spliced into two different mRNAs. In that case, one might expect that two of the exons coding for two of the repeating domains are spliced out to produce the mRNA for the smaller polypeptide.

We initiated our investigations into the structure and biosynthesis of the poly-Ig receptor with the aim of using it as a model system to study intracellular protein traffic. Little is known about the mechanisms whereby membrane proteins are sorted from each other and transported to different locations in the cell⁵⁴. At least some of these steps may require the interaction, directly or indirectly, between the membrane protein and other proteins, for example clathrin. In one simplistic model, the cytoplasmic tail of the transmembrane protein would interact with a cytoplasmic protein. For a membrane protein merely to reach the cell surface the cytoplasmic tail can be as short as two residues, as is the case with membrane-bound IgM⁵⁵. Recent studies on a variety of other membrane proteins that have had their cytoplasmic tails removed or modified by recombinant DNA techniques have yielded conflicting results on whether the tail is necessary for transport to the cell surface^{55–57}. The poly-Ig receptor, on the other hand, has a relatively large cytoplasmic tail of 103 residues. This may reflect the ability of the cytoplasmic tail of the poly-Ig receptor to interact with a number of proteins during its complex transcellular pathway. This pathway requires several sorting steps. After endocytosis via coated pits, the poly-Ig receptor must be sorted away from other proteins that occupy the same initial endosome, but take a different pathway, for example, the asialoglycoprotein receptor⁵⁸. In addition, the poly-Ig receptor is directed first to the basolateral and then to the apical surface of a polarized cell. Having cloned and sequenced the poly-Ig receptor, it may now be possible to determine which portion(s) of the molecule are involved in transporting the protein around the cell and perhaps what other proteins it interacts with along the way.

This work was supported by NIH grant GM-27155. We thank Dr Jacob V. Maizel for his advice and generous provision of computer facilities. We also thank Drs S. Bock, J. Boeke, S. Fallon, A. Hiatt, B. Kay, J. Kaput and K. Krauter for advice on cloning and sequencing, and J. Owens for assistance in the computer-assisted sequence analysis. G. Nimic provided valuable assistance in the preparation of the figures and manuscript.

Received 11 November 1983; accepted 5 January 1984.

1. Brandtzaeg, P. *Clin. exp. Immun.* **44**, 221–232 (1981).
2. Mostov, K. E., Kraehenbuhl, J.-P. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7257–7261 (1980).
3. Kühn, L. C. & Kraehenbuhl, J.-P. *J. biol. Chem.* **256**, 12490–12495 (1981).

4. Mullock, B. M., Hinton, R. H., Dobrota, M., Peppard, J. & Orlans, E. *Biochim. biophys. Acta* **587**, 381–391 (1979).
5. Renston, R. H., Jones, A. L., Christiansen, W. D., Hradek, G. T. & Underdown, B. J. *Science* **208**, 1276–1278 (1980).
6. Nagura, H., Nakane, P. & Brown, W. R. *J. Immun.* **123**, 2359–2368 (1979).
7. Simionescu, N. *Adv. Inflammation Res.* **1**, 64–70 (1979).
8. Mostov, K. E. & Blobel, G. *J. Biol. Chem.* **257**, 11816–11821 (1982).
9. Szul, E. S., Howell, K. & Palade, G. E. *J. Cell Biol.* **97**, 1582–1591 (1983).
10. Tomasi, T. B., Tan, E. M., Solomon, A. & Prendergast, R. A. *J. exp. Med.* **212**, 101–124 (1965).
11. South, M. A., Cooper, M. D., Wollheim, F. A., Hong, R. & Good, R. A. *J. exp. Med.* **123**, 615–627 (1966).
12. Brandtzaeg, P. *J. Immun.* **112**, 1553–1559 (1974).
13. Orlans, E., Peppard, J., Fry, J. F., Hinton, R. H. & Mullock, B. M. *J. exp. Med.* **150**, 1577–1581 (1979).
14. Crago, S. S., Kulhavy, R., Prince, S. J. & Mestecky, J. *J. exp. Med.* **147**, 1832–1837 (1978).
15. Socken, D. J., Jeejeebhoy, K. N., Bazin, H. & Underdown, B. J. *J. exp. Med.* **150**, 1538–1548 (1979).
16. Jackson, G. D. F., Lemaitre-Coelho, L., Vaerman, J. P., Bazin, H. & Beckers, A. *Eur. J. Immun.* **8**, 123–126 (1978).
17. Sanger, F., Nicklen, S. & Coulson, A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
18. Kühn, L. C. *et al. J. Biol. Chem.* **258**, 6653–6659 (1983).
19. Blobel, G. *et al. Symp. Soc. exp. Biol.* **33**, 9–36 (1979).
20. Vite, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
21. Lachlin, A. D. *J. molec. Biol.* **61**, 409–424 (1971).
22. Izel, J. V. & Lenk, R. P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7665–7669 (1981).
23. Dayhoff, M. D., Barker, W. C. & Hunt, L. P. *Meth. Enzym.* **91**, 524–545 (1983).
24. Edelman, G. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **73**, 78–85 (1969).
25. Wilbur, W. J. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 726–730 (1983).
26. Dayhoff, M. O. *et al. Atlas of Protein Sequence and Structure*, Version 6 (1983).
27. Dayhoff, M. O. *et al. Nucleic Acid Sequence Database* (National Biomedical Research Foundation, Washington DC, 1981).
28. Eulitz, M., Gotze, D. & Hilschmann, N. Z. *Physiol. Chem.* **353**, 487–491 (1972).
29. Rudikoff, S. & Potter, M. *J. Immun.* **127**, 191–194 (1981).
30. Campbell, D. G., Gagnon, J., Reid, K. B. M. & Williams, A. F. *Biochem. J.* **195**, 15–30 (1981).
31. Putnam, F. W., Liu, Y.-S. V., & Low, T. L. K. *J. Biol. Chem.* **254**, 2865–2874 (1979).
32. Orr, H. T., Lopez de Castro, J. A., Lancet, D. & Strominger, J. L. *Biochemistry* **18**, 5711–5720 (1979).
33. Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. *Sequences of Proteins of Immunological Interest* (U.S. Department of Health and Human Services, 1983).
34. Ohno, S., Matsunaga, T., Wallace, R. B. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1999–2002 (1982).
35. Edmondson, A. B., Ely, K. R., Abola, E. E., Schiffer, M. & Panagiotopoulos, N. *Biochemistry* **14**, 3953–3961 (1975).
36. Grasbeck, R. & Kouvonen, I. *Trends biochem. Sci.* **8**, 203–205 (1983).
37. Rossi, G., Newman, S. A. & Metzger, H. *J. Biol. Chem.* **252**, 704–711 (1977).
38. Mellman, I. S. & Unkeless, J. C. *J. exp. Med.* **152**, 1048–1069 (1980).
39. Johnson, P. M. *Placenta* **2**, 355–370 (1981).
40. Rodewald, R. *J. Cell Biol.* **85**, 18–32 (1979).
41. Czech, M. P., Massague, J. & Pilch, P. F. *Trends biochem. Sci.* **6**, 222–225 (1981).
42. Strominger, J. L. *et al. Topics dev. Biol.* **14**, 97–113 (1980).
43. Kronenberg, M., Kraig, E. & Hood, L. *Cell* **34**, 327–329 (1983).
44. Williams, A. F. & Gagnon, J. *Science* **216**, 696–703 (1982).
45. Eskeland, T. & Brandtzaeg, P. *Immunochemistry* **11**, 161–163 (1974).
46. Knight, K. L., Vetter, M. L. & Malek, T. R. *J. Immun.* **115**, 595–598 (1975).
47. Cunningham-Rundles, C. & Lamm, M. E. *J. Biol. Chem.* **250**, 1987–1991 (1975).
48. Pardo, A. G., Lamm, M. E., Plout, A. G. & Frangione, B. *Molec. Immun.* **16**, 477–482 (1979).
49. Lindh, E. *J. Immun.* **14**, 284–286 (1975).
50. Jerry, L. M., Kunkel, H. G., & Adam, L. *J. Immun.* **109**, 275–283 (1972).
51. Clarke, S. & Harrison, L. C. *J. Biol. Chem.* **258**, 11434–11437 (1983).
52. Knight, K. L., Rosenzweig, M., Lichter, E. A. & Hanly, W. C. *J. Immun.* **112**, 877–882 (1974).
53. Rogers, J. *et al. Cell* **20**, 303–312 (1980).
54. Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1496–1500 (1980).
55. Zuniga, M. C. *et al. Cell* **34**, 535–544 (1983).
56. Rose, J. K. & Bergmann, J. E. *Cell* **34**, 513–524 (1983).
57. Garoff, H., Kondor-Koch, C., Pettersson, R. & Burke, B. *J. Cell Biol.* **97**, 652–658 (1983).
58. Courtoy, P. J., Quintart, J., Limet, J. N., DeRoe, C. & Bandhuin, P. *J. Cell Biol.* **95**, 425a (1982).
59. Norgard, M. V., Tocci, M. J. & Monahan, J. J. *J. Biol. Chem.* **255**, 7665–7672 (1980).
60. Birnboim, H. C. & Doly, J. *Nucleic Acids Res.* **7**, 1513–1523 (1979).
61. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
62. Gruenstein, M. & Hogness, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961–3965 (1975).
63. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–242 (1977).
64. Messing, J., Crea, R. & Seeburg, P. *Nucleic Acids Res.* **9**, 309–321 (1980).
65. Jung, A., Sippel, A. E., Crez, M. & Schutz, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5759–5763 (1980).
66. Hagenbuchle, O., Borez, R. & Young, R. A. *Cell* **21**, 179–187 (1980).

LETTERS TO NATURE

Rings and wiggles in Hercules A

J. W. Dreher

Department of Physics, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA

Eric D. Feigelson

Department of Astronomy, Pennsylvania State University,
University Park, Pennsylvania 16802, USA and Center for Space
Research, Massachusetts Institute of Technology, Cambridge,
Pennsylvania 02139, USA

Hercules A (3C348) is the fourth brightest extragalactic radio source in the sky at low frequencies. It is optically identified with a 18.5 mag galaxy with redshift $z = 0.154$ located in a faint cluster^{1,2}. Although discovered as early as 1948³, and extensively studied during the 1960s, it was not mapped with high-resolution aperture-synthesis radio telescopes because of its low declination. We report here the first results from extensive observations of Her A made with the Very Large Array (VLA)⁴ of the National Radio Astronomy Observatory. We find that Her A possesses an unusual jet-dominated morphology instead of the Cygnus-A-like morphology expected for such a powerful radio source. Its two jets are quite different in appearance; one appears to be a continuous twisting jet, while the other suggests repeated ejection of individual plasmons by the central object. The presence of such different morphologies in the same source raises interesting questions concerning the generation and evolution of extended extragalactic radio sources. Variations in the central engine provide the simplest interpretation for some of the features of this object; the role of environmental effects is less clear.

VLA observations at 5 GHz were made between January and September 1982. The A, B, and C array configurations were used providing projected baselines ranging from 65 m to 35 km.

Figure 1 shows a map with a 0.5 arc s beam (FWHM) made from a concatenation of the visibilities from all three configurations. This combination of data from different configurations reveals complex, low surface brightness features over a large range of angular scales. The data were calibrated, mapped and cleaned in the usual fashion, after which final adjustments to the complex gains of each of the 27 antennae as a function of time were derived using the self-calibration algorithm⁵. Figure 2 is a low-resolution map produced in a similar manner using only the C configuration data. A deep X-ray image of Her A was also obtained with the Imaging Proportional Counter on the Einstein X-ray Observatory (HEAO-B satellite). A detailed presentation of all of these data, including radio polarization and spectral index information, will be given elsewhere (E.D.F. and J.W.D., in preparation).

Her A has many interesting features. First, it is one of the most luminous radio galaxies known. With $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $q_0 = 0$, its power density at 5 GHz $P_{5,000}$ is $6 \times 10^{26} \text{ W Hz}^{-1}$, or about a quarter that of Cyg A. Because of its large size (460 kpc at greatest extent), its total energy content is about the same as that of Cyg A, $\sim 3 \times 10^{60}$ ergs. Second, despite its high luminosity, it has the morphology of a low luminosity source. Sources of high luminosity ($P_{5,000} > 10^{25} \text{ W Hz}^{-1}$) are generally edge-brightened with conspicuous hotspots at the outer edges of their radio lobes (Fanaroff-Riley⁶ Class II morphology). Lower luminosity sources are edge-darkened (F-R Class I) and frequently have jets emerging from both sides of the core⁷. Third, the eastern jet contributes a substantial fraction of the total luminosity, making it the highest luminosity jet yet found associated with an elliptical galaxy. More luminous jets have been found only in extended radio quasars which are far more distant than Her A. Fourth, the western jet exhibits several remarkable ring-like structures. Although rings have been noticed in a few other sources (3C219⁸, 3C310⁹, and recently Cyg A¹⁰), those in Her A are much more distinct and form a linear sequence leading from the core to the lobes. We discuss these rings in more detail below. Fifth, the low resolution map (Fig. 2) reveals smooth

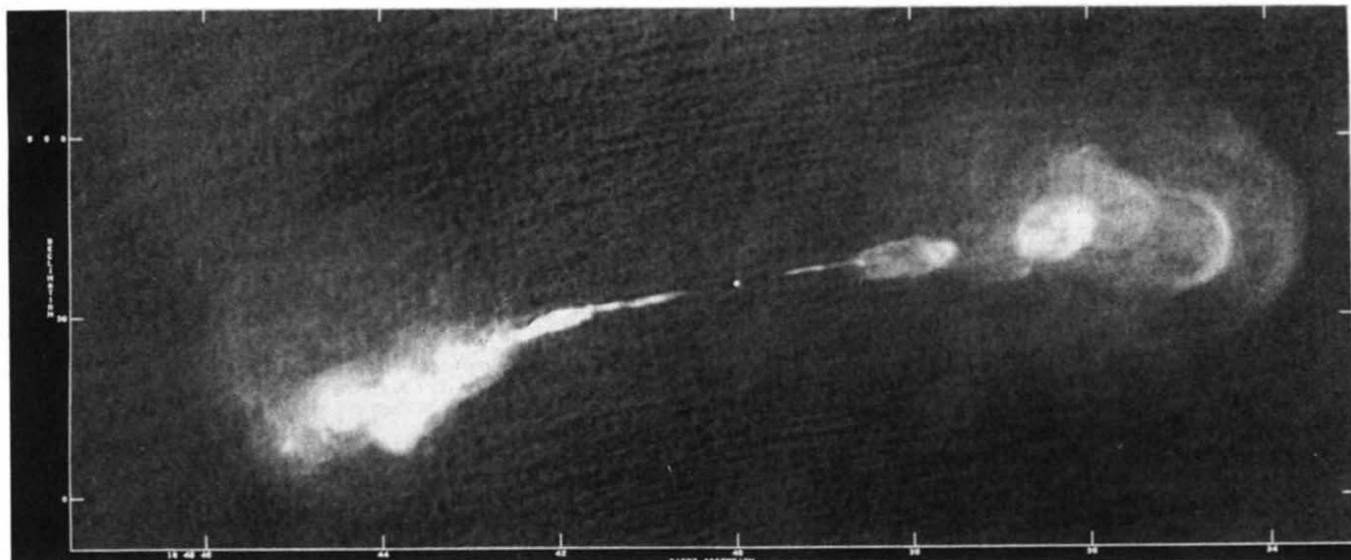


Fig. 1 A 5-GHz VLA map of Her A with a 0.5 arc s beam. The map is shown as a grey-scale image, with the range of shades indicated on the bar below the image corresponding to radio brightnesses from -0.05 mJy per beam to 5 mJy per beam. Brightnesses between 5 mJy per beam and 22 mJy per beam, the maximum on this map, are all shown as white. The source is 3 arc min across.

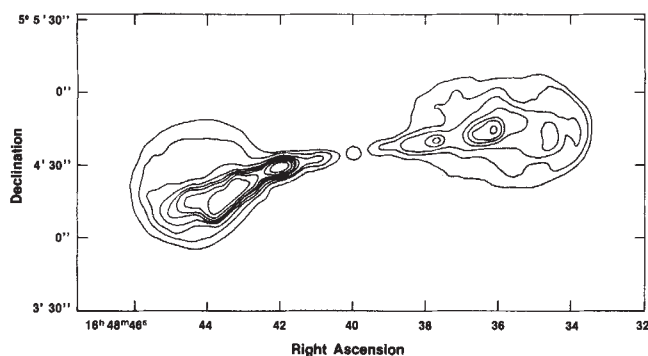


Fig. 2 A low-resolution map (beam 4.2 arc s, observations at 5 GHz) with contours at 0.5, 2, 5, 10, 15, 20, 30, 50, and 90% of the maximum of 511 mJy per beam.

symmetric lobes enveloping the two jets. A much fainter bridge spanning these two lobes and surrounding the core can also be seen on preliminary 1.4 GHz VLA maps. Finally, the Einstein Observatory X-ray image of Her A shows an extended source with $L_x = 1.5 \times 10^{44} \text{ erg s}^{-1}$ (0.2–4 keV) and diameter ≥ 1.0 Mpc. These parameters are typical¹¹ of a cluster of galaxies with Abell richness class 0 to 1.

On the high resolution map (Fig. 1) it can be seen that the western jet in fact consists of an inner jet followed by what appear to be at least three full or partial ring-like features. The inner jet begins 9 arc s from the radio core, extends out to a bright spot on the inner edge of the first ring 21 arc s from the core and may even continue through this loop to terminate on its further side. The bright regions of this inner jet are at most slightly resolved with our 0.5 arc s beam; the opening angle is $< 3^\circ$ (half-width to half-maximum as seen from the core). The rings are brightest on their leading or outward edges. The outermost, in particular, is visible only on this edge. (Note that the furthest arc visible on Fig. 1, at 16 h 48 min 34 s, is the outer edge of the western lobe and not necessarily part of a distinct ring.) The transverse widths of the rings increase linearly with distance from the core. The outer two arcs are $\sim 50\%$ linearly polarized. The E vectors of the polarized radio intensity are generally normal to these arcs, suggesting that the magnetic fields lie along them. The inner loops are less polarized and have a complex polarization distribution.

The apparent sequence of discrete rings suggests successive episodes of ejection from the active nucleus. The limb-brightening could then indicate that relativistic particles are reacceler-

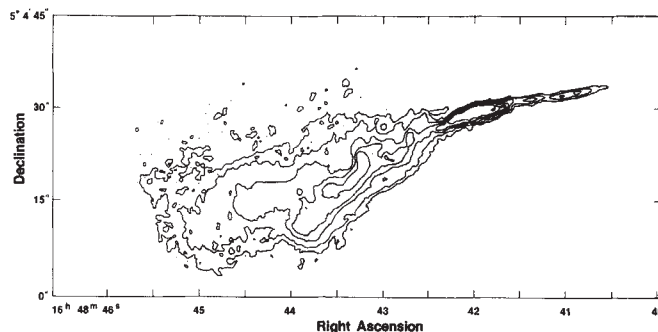


Fig. 3 A full-resolution map of the eastern jet. The contours are at intervals of 2, 4, 10, 20, 40, and 90% of the peak brightness of 22.0 mJy per beam. Beam is 0.5 arc s, observations at 5 GHz.

ated at the plasmon surface, perhaps due to interactions with the ambient medium^{12,13}. Reacceleration and, possibly, field amplification would be necessary to balance the adiabatic losses that must occur in this picture. The change from the inner elongated ring to the rounder outer rings, and the small knot at the base of the first ring might both be caused by differences in the velocities of the ejecta. These velocity differences might reflect either deceleration by the external medium, or variable velocities of outflow at the central engine as suggested by Rees¹⁴. However, alternative explanations for the rings are certainly also possible, such as instabilities that act to 'pinch off' segments of a steady beam or explosive encounters between an invisible beam and ambient clouds^{15,16}.

The brighter eastern jet has an appearance very different from that of the western jet. Figure 3 is a high-resolution contour map of this jet. It can be traced continuously for at least 55 arc s from the core and exhibits three distinct regimes over this length. It first appears 9 arc s from the core as a narrow jet with three intensity maxima. Its opening angle is $< 3^\circ$. Next, beginning 22 arc s from the core the jet suddenly brightens, makes a small kink, and widens to ~ 2 arc s FWHM. Within this bright region several small knots are present. As this region fades, the jet appears to narrow slightly. Here, ~ 40 arc s from the core, it enters a regime in which it widens and undergoes back-and-forth lateral displacements of increasing amplitude, or 'wiggles'. Ultimately the jet is lost in a region of fainter complex structure. The eastern jet is 30 to 40% polarized over its entire length, with the suggested magnetic field parallel to the jet and following the wiggles.

Considered by itself, the eastern jet is similar to those in a number of other well-studied sources. The jets in NGC315¹⁷ and NGC6251¹⁸ also show periods of widening and recollimation, and the development of increasingly prominent wiggles is seen in M87¹⁹ and 3C449²⁰, among others. Two classes of models have been discussed to explain these wiggles: orbital²¹ or precessional²² motions of the central engine, and instabilities, such as Kelvin–Helmholtz^{23–26} or ‘firehose’^{27,28}, in the outflowing material. However, the presence of the relatively straight inner jets seems to rule out simple models of the first class, and we argue below that instabilities *alone* cannot explain this jet either.

Are the jet morphologies determined by variations of the central engine, as suggested by the appearance of the western jet, or by instabilities, as suggested by the appearance of the eastern jet? A clue to the resolution of this ‘nature’ versus ‘nurture’ dichotomy can be found in the underlying symmetry of the source. The low resolution map of Fig. 2 clearly reveals that on scales of >10 arc s the principal features of the eastern and western jets are symmetrically located about the core: the narrow inner jets on each side are very similar (at both resolutions), the brightest part of the eastern jet and the first ring in the west both lie between 21 and 38 arc s from the core, and, finally, the bright wiggle in the east and the second ring in the west coincide at 45–60 arc s from the core. This correspondence between major features on both sides of the source implies that the basic structure of both jets is determined by a common factor, of which variations in the central engine seem most plausible. It seems unlikely that models based on interactions of the jets or beams with their environments will be able to account for the overall symmetry. However, it remains true that the detailed structures of the two jets are strikingly different. Thus it may be necessary to devise a model for this source in which the radio morphology is regulated both by its origin in a variable central engine and by differences in the development of each jet as it propagates outward. Other radio jets associated with central dominant galaxies (such as 3C465²⁹) also exhibit unusual morphologies as they propagate through the external medium. These differences could either reflect an asymmetry in the environment or in the central engine³⁰.

We thank Roger Blandford, Alan Bridle and Phil Morrison for helpful comments, and George Clark for his many contributions to this project. This research was partially supported by NASA contract NAS8-30752 and NSF grant AST-82-10996. The National Radio Astronomy Observatory is operated by Associated Universities, Inc. under contract with the NSF.

Received 30 September; accepted 14 December 1983.

- Williams, P. J., Dewhurst, D. W. & Leslie, P. R. *Observatory* **81**, 64–67 (1961).
- Greenstein, J. L. *Astrophys. J.* **135**, 679–683 (1962).
- Bolton, J. G. *Nature* **162**, 141–143 (1948).
- Thompson, A. R., Clark, B. G., Wade, C. M. & Napier, P. J. *Astrophys. J. Suppl.* **44**, 151–167 (1980).
- Schwab, F. *Proc. Soc. Photo-Optical Instr. Engin.* **231**, 18–23 (1980).
- Fanaroff, B. L. & Riley, J. M. *Mon. Not. R. astr. Soc.* **167**, 31P–35P (1974).
- Bridle, A. H. *IAU Symp.* No. 97, 121–128 (1982).
- Perley, R. A., Bridle, A. H., Willis, A. G. & Fomalont, E. B. *Astr. J.* **85**, 499–506 (1980).
- van Breugel, W. & Fomalont, E. B. *Astrophys. J.* (submitted).
- Perley, R. A., Cowan, J. & Dreher, J. W. (in preparation).
- Kriss, G. A., Cioffi, D. F. & Canizares, C. R. *Astrophys. J.* **272**, 439–448 (1983).
- Christiansen, W. A., Rolison, G. G. & Scott, J. S. *Astrophys. J.* **234**, 456–465 (1979).
- Eilek, J. A. *Astrophys. J.* **254**, 472–482 (1982).
- Rees, M. J. *Mon. Not. R. astr. Soc.* **184**, 61P–65P (1978).
- Blandford, R. D. & Konigl, A. *Astrophys. Lett.* **20**, 15–21 (1980).
- Morrison, P., Roberts, D. & Sadun, A. (submitted).
- Willis, A. G., Strom, R. G., Bridle, A. H. & Fomalont, E. B. *Astr. Astrophys.* **95**, 250–265 (1981).
- Perley, R. A., Bridle, A. H. & Willis, A. G. *Astrophys. J. Suppl.* (in the press).
- Owen, F. N., Hardee, P. E. & Bignell, R. C. *Astrophys. J. Lett.* **238**, L11–L15 (1980).
- Perley, R. A., Willis, A. G. & Scott, J. S. *Nature* **281**, 437–442 (1979).
- Blandford, R. D. & Icke, V. *Mon. Not. R. astr. Soc.* **185**, 527–538 (1978).
- Gower, A. C., Gregory, P. C., Hutchings, J. B. & Unruh, W. G. *Astrophys. J.* **262**, 478–496 (1982).
- Ferrari, A., Massaglia, S. & Trussomi, E. *Mon. Not. R. astr. Soc.* **198**, 1065–1080 (1982).
- Hardee, P. E. *Astrophys. J.* **269**, 94–101 (1983).
- Ray, T. R. *Mon. Not. R. astr. Soc.* **198**, 617–625 (1982).
- Norman, M. L., Smarr, L., Winkler, K.-H. & Smith, M. D. *Astr. Astrophys.* **113**, 285–302 (1982).
- Benford, G. *Mon. Not. R. astr. Soc.* **183**, 29–48 (1978).
- Eilek, J. A., Burns, J. O., Owen, F. N. & O’Dea, C. P. *Astrophys. J.* (in the press).
- Icke, V. *Astrophys. J.* **265**, 648–663 (1983).

An isolated baroclinic eddy as a laboratory analogue of the Great Red Spot on Jupiter

P. L. Read & R. Hide

Geophysical Fluid Dynamics Laboratory, Meteorological Office, Bracknell, Berkshire RG12 2SZ, UK

We have recently presented evidence¹ supporting the hypothesis² that the long-lived large oval atmospheric eddies on Jupiter and Saturn, including Jupiter’s anticyclonic Great Red Spot (GRS) and White Ovals and the cyclonic ‘barges’, are manifestations of ‘slantwise’ or ‘sloping’ convection in a rotating fluid, implying that they are involved in the horizontal transport of heat towards or away from the edges of the atmospheric zones or belts in which they occur, and that their kinetic energy derives directly from the action of gravity on the density field associated with horizontal gradients of temperature. These eddies would then be dynamically similar to the highly stable closed ‘baroclinic’ eddies produced in certain laboratory experiments on thermal convection in a rotating fluid subject to internal heating or cooling^{1,3}. We now present laboratory findings that bear on the interpretation of the isolated nature of the GRS, including the crucial demonstration that a single intense stable baroclinic disturbance that is strongly localized in azimuth can form readily when the impressed conditions are close to the transition from axisymmetric to non-axisymmetric flow.

The laboratory experiments described in refs 1–3 form part of a much wider study of thermal convection in a rotating fluid annulus^{4–8}. The working fluid of mean density ρ_0 occupies an annular container with two vertical, coaxial, thermally conducting cylindrical sidewalls, and thermally insulating endwalls. The whole apparatus is rotated steadily about its vertical axis of symmetry. The character of the flow obtained in the rotating annulus depends on the external conditions^{5–8}, which can be specified in terms of several dimensionless parameters involving the axial and transverse dimensions of the apparatus, d and $(b-a)$ respectively, the rotation rate Ω , acceleration due to gravity g (typically very much greater than $\Omega^2 b$) the maximum density contrast $\Delta\rho$ associated with the axisymmetric impressed temperature field, and the kinematic viscosity ν and other physical properties of the working fluid. The most important of these parameters are:

$$\Theta \equiv \frac{g d \Delta\rho}{\Omega^2 \rho_0 (b-a)^2} \quad (1)$$

$$\mathcal{T} \equiv \frac{4\Omega^2 (b-a)^5}{\nu^2 d} \quad (2)$$

The flow itself strongly modifies the horizontal density field associated with the impressed differential heating and cooling, thereby reducing the horizontal density gradients in the main body of the fluid, away from sidewall and endwall boundary layers. It also produces vertical density gradients which are typically strongly dependent on the upward vertical co-ordinate z , but comparatively weakly dependent on the horizontal co-ordinates (r, θ) . Associated with these vertical gradients can be defined a local ‘Brunt–Väisälä’ frequency

$$N \equiv \left[-\frac{g}{\rho_0} \frac{\partial \rho}{\partial z} \right]^{1/2} \quad (3)$$

the average value of which is typically⁹ $\sim (0.8g\Delta\rho/\rho_0 d)^{1/2}$. When $\mathcal{T}$ is very much greater than 2×10^5 , the character of the flow is found^{3,5–7} to depend mainly on Θ . At values of Θ exceeding a certain critical value Θ_R of order unity (whose exact value depends on details of mechanical and thermal boundary conditions and with which⁹ there is an associated critical average value of N), the flow is axisymmetric, that is, its properties are

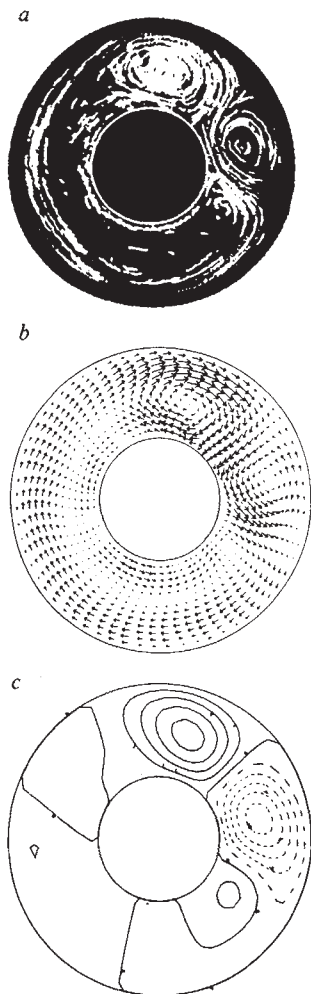


Fig. 1 The instantaneous upper level flow field ($z/d = 0.7$ where $z = d$ corresponds to the upper surface) showing a steady, isolated baroclinic eddy in a rotating fluid subject to internal heating and sidewall cooling. *a*, Streak photograph (10-s streaks) from the laboratory experiment. *b*, Horizontal velocity vectors derived from velocity measurements using streak images similar to *a* via the method of Jonas and Kent¹⁰; the maximum velocity indicated is 0.4 cm s^{-1} . *c*, Eddy stream function, with the azimuthal mean component removed, obtained from the interpolated velocity field shown in *b*. Negative contours are dashed.

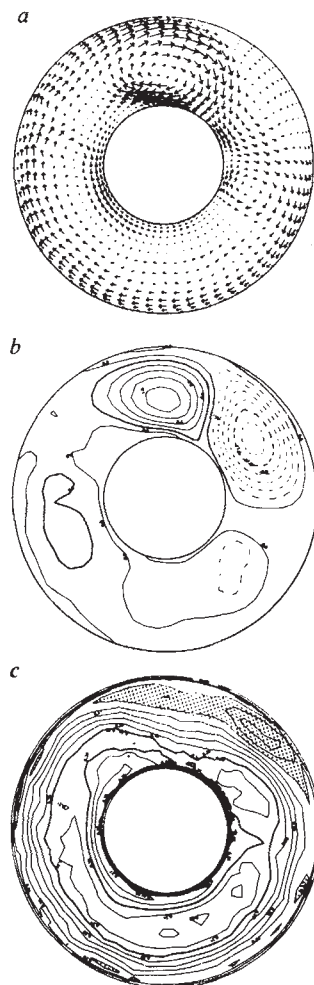


Fig. 2 The instantaneous upper level ($z/d = 0.7$) flow field with an isolated baroclinic eddy in a rotating fluid subject to internal heating and sidewall cooling, obtained from a numerical simulation after 20 min of simulated time since the introduction of the perturbation. *a*, Horizontal velocity vectors (maximum velocity indicated is 0.83 cm s^{-1}). *b*, Eddy kinetic pressure field (that is, with azimuthal mean component removed), with contour interval $0.2 \text{ cm}^2 \text{ s}^{-2}$ and negative contours are dashed. *c*, Vertically-averaged square of the Brunt-Väisälä frequency, N^2 (a measure of the local static stability) averaged over the interval $0.07 \leq z/d \leq 0.87$, with a contour interval of 0.01 s^{-2} . The heavy contour is at 0.30 s^{-2} and the hatched region indicates where $N^2 < 0.25 \text{ s}^{-2}$.

independent of the azimuthal coordinate θ . Regular nonaxisymmetric flows (that is, spatially periodic with dominant azimuthal wavenumber $m \neq 0$, and either steady or temporally periodic) occur within the range $\Theta_R > \Theta > \Theta_1$, where Θ_1 is a further critical value of Θ . When $\Theta < \Theta_1$ irregular, aperiodic nonaxisymmetric flows occur, owing^{5,7} to the onset of barotropic instability associated with the comparatively small azimuthal scale (proportional to m^{-1}) of the baroclinic waves. Both regular and irregular nonaxisymmetric flows exhibit fully-developed 'baroclinic' disturbances, in which the kinetic energy of the flow is generated and maintained against viscous dissipation by 'slantwise' convection from the potential energy associated with the density field resulting from the impressed heating and cooling.

The detailed form of the flow pattern in the regular nonaxisymmetric regime depends *inter alia* on the variation with distance r from the axis of symmetry of the impressed temperature field^{3,5-8}. In the particular configurations considered in ref. 1, heat is introduced (or extracted) internally and extracted (or introduced) via both sidewalls, with the corresponding r variation of impressed temperature showing a maximum (or minimum) near mid-radius. The upper level flow pattern largely consists of m identical closed, compact, equally-spaced oval eddies. Near the edge of each eddy, the relative flow is concentrated into a 'peripheral jet-stream' circulating in the same sense as the shear of the background azimuthal flow, which is anticyclonic (cyclonic) near the upper surface when the system is subject to internal heating (cooling) and cooling (heating) at both sidewalls.

The flows described in the present paper were generated using the same apparatus, numerical model and configurations of heating and cooling as before¹, with attention being concentrated on conditions just within the regular nonaxisymmetric flow regime, where Θ is only slightly less than Θ_R , at compara-

tively large values of $\mathcal{T}$ (in excess of 10^7). The upper level nonaxisymmetric flow in a typical laboratory experiment near $\Theta = \Theta_R$ is dominated by a steady isolated pair of baroclinic eddies. In Fig. 1, panel *a* shows a streak photograph of the flow, and *b* and *c* are derived from the velocity field, obtained using velocities measured from streak images similar to Fig. 1*a* by means of a novel automatic technique¹⁰. The 'eddy' stream function shown in Fig. 1*c* is approximately equivalent to the eddy pressure field, assuming the flow to be geostrophic, and was obtained by integrating the above-mentioned velocity field and subtracting the azimuthal mean component. Figure 1*b* shows the flow pattern at this upper level, where the shear of the background flow is anticyclonic. The flow is clearly dominated by a single large anticyclonic eddy accompanied by a weaker cyclonic feature, together spanning no more than about 100° in azimuth. Elsewhere the flow is virtually independent of θ . The strength of the cyclonic feature in the flow varies with height, dominating the flow at low levels where the shear of the background flow is also cyclonic. As the transitional value Θ_R was approached more closely than the flow in Fig. 1, the upper level cyclonic feature was observed to decrease in strength until it manifested itself as no more than a slight cyclonic curvature of the streamlines near the much more prominent anticyclonic eddy. The eddy stream function in Fig. 1*c* emphasizes the localized nature of the disturbance, which appears as a strongly modulated pulse of waves with a wavelength of about 120° in azimuth. Once established, the pattern is quite steady, apart from a slow steady azimuthal drift relative to the walls of the apparatus at a rate which depends on a variety of factors (for example, endwall slope). An $m = 2$ flow pattern, symmetric about a diameter, was found to occur under the same external conditions as those prevailing for an isolated eddy similar to Fig. 1. This is in keeping with previous findings^{3,5-7} that such

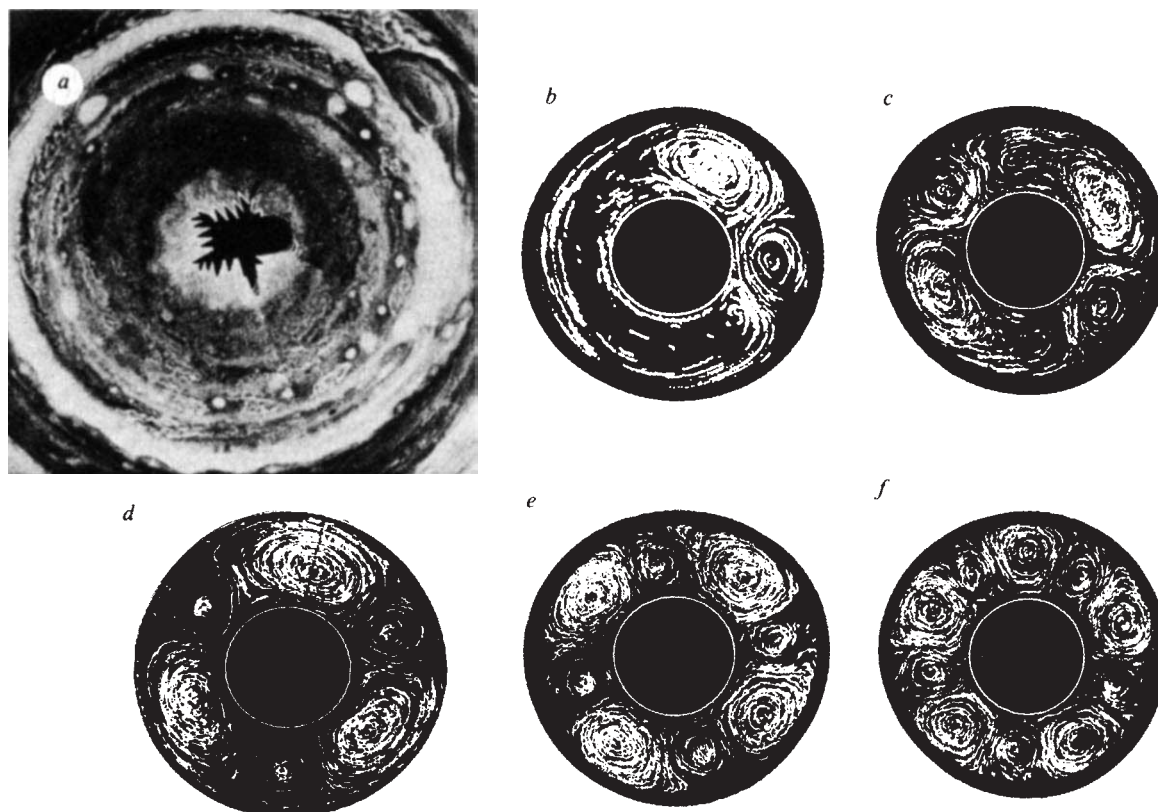


Fig. 3 *a*, A mosaic of Voyager 1 images of Jupiter, projected to show the planet as seen from below the south pole. Latitude lines are concentric circles centred on the pole, which lies within the black jagged area (due to missing data) at the middle of the frame. The long-lived anticyclonic atmospheric eddies appear (apart from the GRS top right) as trains of white oval spots on lines of constant latitude. *b–f*, Upper level streak photographs, obtained in the laboratory at various values of Θ , showing steady regular baroclinic eddies in a rotating fluid subject to internal heating and sidewall cooling. *b*, Isolated eddy (equivalent to $m = 1$, compare Fig. 1*a*), $\Theta = 2.92$; *c*, $m = 2$, $\Theta = 2.84$; *d*, $m = 3$, $\Theta = 2.16$; *e*, $m = 4$, $\Theta = 0.72$; *f*, $m = 5$, $\Theta = 0.48$. The mean radius of the planet Jupiter is 6.97×10^7 m, nearly 10^9 times the outer radius of the annular convection chamber (10^{-1} m), but the appropriate dimensionless similarity parameters are comparable.

systems are 'intransitive' in the sense that m is not uniquely determined by Θ and $\mathcal{T}$.

More detailed data on the structure and energetics of the isolated eddy were obtained from our numerical model^{1,11} as follows. A two-dimensional (axisymmetric) version of the model was integrated to a steady state under similar external conditions to one of the experiments. Then, a localized perturbation of azimuthal width equal to one grid point was introduced into the temperature field, and its subsequent evolution was followed using the full three-dimensional model. Figure 2 shows some representative fields after 20 min of simulated time following the introduction of the perturbation, when the flow was settling down to its final state. The velocity field near the level sampled in the laboratory (see Fig. 1) is shown in Fig. 2*a*, and exhibits characteristics similar to those of the laboratory flows. The flow pattern is dominated by a single large anticyclonic feature accompanied by a weaker cyclonic circulation. The eddy kinetic pressure field in Fig. 2*b* confirms the conclusion from the laboratory data (see Fig. 1*c*) that the eddy pair spans little more than 100° in azimuth.

An indication of the mechanism by which the eddy might develop as a localized feature is provided by Fig. 2*c*. This shows the instantaneous distribution in (r, θ) of N^2 , the vertically-averaged square of N (see equation (3) above). Marked azimuthal variations in N^2 are apparent, even though the distribution of impressed heating and cooling is axisymmetric, with the isolated eddy occupying the range in azimuth where N^2 reaches a minimum. The entire pattern of N^2 drifts around the apparatus with the eddy. The azimuthal variations in N^2 are produced by the baroclinic disturbance itself, the nonlinear effects of advection serving to reduce the overall static stability such that it is only in the vicinity of the eddy that the local value of N^2 is sufficiently small for sloping convection to occur. These results demonstrate the effect of nonlinear interactions between

the eddy and its surroundings, which tend to sharpen its distribution in azimuth. The crucial new discovery, however, is that when Θ is only slightly less than Θ_R the disturbance equilibrates to a highly localized form, even though the boundary conditions and heating are axisymmetric. Further studies are clearly required to examine the details of this mechanism, and to understand why the initial disturbance remains localized in azimuth and does not extend around the entire annulus, as occurs in the model under conditions when Θ is significantly less than Θ_R .

The results of ref. 1 and the present work demonstrate convincingly that fully-developed 'slantwise convection' can account for many of the observed properties of the long-lived eddies in the atmospheres of Jupiter and Saturn. Accordingly, the sense of circulation and oval form of the atmospheric eddies result from the non-monotonic latitudinal variation of the background thermal field, the associated zonal flow having strong horizontal (as well as vertical) shear. The compact nature of the eddies, their peripheral jet-streams and associated distribution of vertical motion, and their longevity, are characteristic properties of baroclinic eddies over the wide range of parameter space occupied by the regular nonaxisymmetric regime. The isolated nature of the most striking of all the atmospheric eddies, the GRS, is evidently characteristic of the regular nonaxisymmetric flow close to the transition to axisymmetric flow. The latitudinal scale of the eddies is governed by the spacing of the atmospheric bands of wind and cloud, which we presume here to be produced by other, as yet not fully understood, processes^{12,13}. The variation with latitude of the number of the oval eddies and their longitudinal scale, especially evident on Jupiter (see Fig. 3), most probably reflects variations in the width of the atmospheric bands and in the effective local value of Θ (see equation (3)) where Ω is now the local vertical component of the planet's rotation vector. According to the present work, the value of Θ

at the latitude of the GRS should be close to Θ_R , and N^2 should reach a minimum at the longitude of the GRS. The former result could be consistent with the absence of a large eddy comparable to the GRS at similar northern latitudes on Jupiter, since very small asymmetries in variables such as net heating between the two hemispheres could result in Θ exceeding Θ_R at higher latitudes in the northern hemisphere than in the southern hemisphere. We note also the property of intransitivity revealed by the laboratory experiments (see above), implying that a pattern of flow that includes two or more regularly spaced eddies identical to the GRS may also be possible under the conditions now prevailing at the latitude of the present GRS (compare Fig. 3b and c).

At first sight it might seem remarkable that dynamically similar phenomena can be produced on scales differing by such a large factor, but the theoretical investigation of the dynamical processes that produce baroclinic eddies can account for this similarity^{2,3,5-8,14-17}. It is sufficient to say here that the flow in the main body of the fluid, outside the boundary layers at the top, bottom and sides, is quasigeostrophic and satisfies the so-called 'potential vorticity' equation

$$\frac{D}{Dt} \left[\nabla^2 \Psi + f^2 \frac{\partial}{\partial z} \left(\frac{1}{N^2} \frac{\partial \Psi}{\partial z} \right) \right] - \frac{\alpha}{r} \frac{\partial \Psi}{\partial \theta} = 0 \quad (4)$$

Here Ψ is a stream function related to the vertically-averaged horizontal flow, ∇^2 is the horizontal laplacian operator (so that $\nabla^2 \Psi$ is the average z -component of relative vorticity), D/Dt is the substantial time derivative following the geostrophic motion, and $f = 2\Omega$. The term $-\alpha r^{-1} \partial \Psi / \partial \theta$ represents potential vorticity changes associated with the combined effects of axisymmetric sloping endwalls and variations in the vertical density stratification, where α depends on the values of the endwall slopes and the thermal structure of the flow^{16,17}. The term is roughly analogous to the so-called 'beta' term in the corresponding form of the equation used in dynamical meteorology and oceanography^{14,15}, where α is there taken as the rate of change with latitude of the vertical component of the Earth's rotation vector. An equation of this form governs many types of quasigeostrophic flows, including slantwise convection, in systems ranging from planetary atmospheres and oceans with horizontal scales of thousands (and even tens of thousands) of kilometres down to the laboratory apparatus used in the experiments described in the present paper. This remarkable result can be understood by noticing that the essential balance of terms in equation (4) is such that

$$\left| \frac{D}{Dt} (\nabla^2 \Psi) \right| \sim \left| \frac{D}{Dt} \left\{ f^2 \frac{\partial}{\partial z} \left(\frac{1}{N^2} \frac{\partial \Psi}{\partial z} \right) \right\} \right| \gg \left| \frac{\alpha}{r} \frac{\partial \Psi}{\partial \theta} \right| \quad (5)$$

This implies that $fL/NH \sim 1$ if L is a characteristic horizontal dimension associated with the term $\nabla^2 \Psi$, and H is a characteristic vertical dimension associated with the term $f^2 \partial(N^{-2} \partial \Psi / \partial z) / \partial z$. For example, in the case of the Earth's atmosphere, $L \leq 10^7$ m, $f \sim 10^{-4}$ s⁻¹, $N \sim 10^{-2}$ s⁻¹ and $H \sim 10^4$ m, so that $fL/NH \sim 10$. For typical laboratory annulus experiments $L \leq 10^{-1}$ m, $0 < f < 10$ s⁻¹, $0.1 < N < 1$ s⁻¹ and $3 \times 10^{-2} < H < 3 \times 10^{-1}$ m, enabling a wide range of values of fL/NH to be achieved, from zero up to ~ 300 . This includes values typical of planetary atmospheres. For the south tropical zone of Jupiter, where the GRS occurs, $L \sim 10^7$ m and $f \sim 10^{-4}$ s⁻¹; this implies $NH \sim 10^3$ m s⁻¹ if for quasigeostrophic disturbances we have $fL/NH \sim 1$. This value of NH is consistent with what little is known about the deep vertical structure of Jupiter's atmosphere^{1,12,13}. ($N/f \sim L/H$ is sometimes referred to as the 'Prandtl ratio of scales', and, in special cases where H is a well-defined quantity, $NH/f \sim L$ is termed the 'Rossby radius of deformation'.) The 'alpha' term $-\alpha r^{-1} \partial \Psi / \partial \theta$ in equation (4) can modify slantwise convection quantitatively, by changing the azimuthal and, in extreme cases, radial wavelengths of unstable modes and their rate of propagation relative to the mean flow, and altering the critical values of Θ at which axisymmetric flow gives way to nonaxisymmetric flow^{16,17}. But there

is ample evidence that the 'alpha' term does not preclude the existence of the regular nonaxisymmetric regime of flow that is invoked in the hypothesis discussed in this paper.

We thank Dr P. Hignett and Dr A. A. White for useful discussions and for their comments on the manuscript, and Mr I. M. Armstrong, Mr R. D. Carter and Mr W. D. N. Jackson for technical assistance.

Received 21 June; accepted 8 December 1983.

1. Read, P. L. & Hide, R. *Nature* **302**, 126-129 (1983).
2. Hide, R. *Observatory* **100**, 182-193 (1980); *Met. Mag.* **110**, 335-344 (1981).
3. Hide, R. & Mason, P. J. *Phil. Trans. R. Soc. A* **268**, 201-232 (1970).
4. Ukaji, K. *J. met. Soc. Japan* **57**, 532-547 (1979).
5. Hide, R. & Mason, P. J. *Adv. Phys.* **24**, 47-100 (1975).
6. Fultz D., Long R. R., Owens G. V., Bohan W., Kaylor R. & Weil, J. *Met. Monogr. Amer. met. Soc.* **4**, 1-104 (1959).
7. Hide, R. *Phil. Trans. R. Soc. A* **250**, 441-478 (1958).
8. Hart, J. E. A. *Rev. Fluid Mech.* **11**, 147-172 (1979).
9. Hide, R. *J. atmos. Sci.* **24**, 6-9 (1966); *Phys. Fluids* **10**, 56-68 (1967).
10. Jonas, P. R. & Kent, P. M. *J. Phys. E* **12**, 604-609 (1979).
11. James, I. N., Jonas P. R. & Farnell L. Q. *J. R. met. Soc.* **107**, 51-78 (1981).
12. Gehrels, T. (ed.) *Jupiter: Studies of Interior, Atmosphere, Magnetosphere and Satellites* (University of Arizona Press, Tucson, 1976); *Saturn: Studies of Interior, Atmosphere, Rings and Satellites* (University of Arizona Press, Tucson, 1984).
13. Flasar, F. M. et al. *J. geophys. Res.* **86**, 8759-8767 (1981).
14. Pedlosky J. *Geophysical Fluid Dynamics* (Springer, Berlin, 1979).
15. Gill, A. E. *Atmosphere-Ocean Dynamics* (Academic, London, 1982).
16. Mason, P. J. *Phil. Trans. R. Soc. A* **278**, 397-445 (1975).
17. Hide, R. *The Global Circulation of the Atmosphere* (ed. Corby, G. A.) 196-221 (Royal Meteorological Society, London, 1969).

Polar glaciation and the genesis of ice ages

B. G. Hunt

Australian Numerical Meteorology Research Centre,
PO Box 5089AA, Melbourne, Australia 3001

One of the outstanding, unresolved climatic questions concerns the cause(s) of ice ages over geological time^{1,2}. A recurring theme in many hypotheses²⁻⁴ has been that land at high latitude is a necessary condition, as snow accumulation on such land would accentuate local cooling via its albedo effect. Although the validity of these hypotheses has been queried⁵, they have been applied to the three most recent glacial periods, namely, the Quaternary (present), the Permo-Carboniferous ($\approx 300 \times 10^6$ years ago) and the Ordovician ($\approx 450 \times 10^6$ years ago). The involvement of high-latitude land has now been investigated with a general circulation model of the atmosphere, and the results presented here indicate that sea ice is readily produced in the model at high latitudes regardless of the existence of land. While land, or at least very shallow seas, may be necessary to promote glaciation, it appears that polar glaciation is the norm rather than the exception for the Earth. Thus no particular perturbation to the climatic system may be necessary to explain ice ages; rather the problem may be one of accounting for the intervening periods of warmth. The most promising solution to the latter problem is very large increases in the atmospheric CO₂ content.

Experiments were performed with a large-scale atmospheric model in order to investigate the land/albedo hypothesis. They have been concentrated on the Arctic Ocean as it has been suggested⁶⁻⁸ that its ice cover can be destroyed relatively easily, which could cause it to remain ice free. Previous model studies^{9,10} have shown that substantial atmospheric changes resulted from permanently removing Arctic sea ice, but did not demonstrate whether an ice-free state was sustainable.

The control general circulation model has been described in some detail previously^{11,12}. Essentially it was hemispheric in extent, had 18-levels in the vertical and approximately 1,200 gridpoints per level. It was set up for nondiurnal, annual mean conditions based on specified climatological, time invariant, zonal mean values of water vapour, carbon dioxide, ozone and clouds required for the radiative calculations. Neither mountains

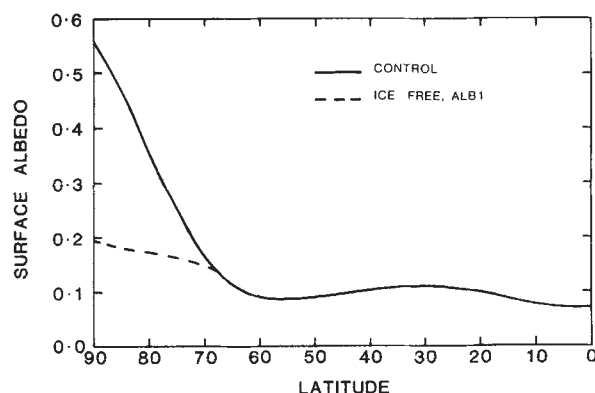


Fig. 1 Zonally averaged, annual mean surface albedos for the control and ALB1 runs.

nor land-sea contrast were included; this appears to be the least arbitrary surface specification in view of the extent of the glacial periods under consideration and the poorly known range of continental dispositions possible. The model surface was effectively a 'swamp', implying zero heat capacity in the surface as regards the calculation of the surface temperature, but that a source of water vapour related to the saturation vapour pressure of the surface temperature existed at the surface. The advantage of such a swamp condition is that it permits the model to compute a surface temperature at all gridpoints compatible with its self-determined climate. The alternative in a model with land-sea contrast is to specify climatologically the sea surface temperature, which would be quite inappropriate for the experiments under consideration here. The model also incorporated the atmospheric hydrological cycle, implying it forecast its own precipitation, but that this forecast was independent of the specified cloud properties used in the radiative transfer code. Despite the simplification employed in the model, it was capable of simulating many of the features of the current climate, including surface temperature¹¹.

A sea-ice model in one of the experiments was a simplified version of that developed by Parkinson and Washington¹³. In this model, starting with the mixed layer ocean, the change in its temperature, T_0 , at any time was determined by the net heat input, Q , due to solar radiation, upwards and downwards long-wave radiation at the surface, sensible heat flux and latent heat flux as determined in the model. Then

$$T_0^{\text{new}} = T_0^{\text{old}} - QDT/(HC)$$

$DT = 10$ min is the time between temperature calculations, $H = 70$ m, the mixed layer ocean depth and $C = 4.19 \times 10^6 \text{ J m}^{-3} \text{ K}^{-1}$, the heat capacity of seawater. If T_0 is below 271.2 K, the freezing point of seawater, 100% of the model grid square was assumed to be ice covered; no snow accumulation on the ice was considered and no movement of the ice was permitted. The initial sea ice depth was:

$$Z = (271.2 - T_0)HC/k_t$$

where $k_t = 3.02 \times 10^8 \text{ J m}^{-3}$ is the heat of fusion of ice. In the presence of sea ice Q is modified to include a heat flux through the ice, F_i , and an oceanic heat flux, F_o , into the ice.

$$F_i = k_c(T_0 - T^*)/Z$$

where $k_c = 2.04 \text{ J m}^{-1} \text{ s}^{-1} \text{ K}^{-1}$ is the thermal conductivity of ice and T^* is the surface temperature of the ice. An oceanic heat flux of $10 \text{ J m}^{-2} \text{ s}^{-1}$ was assumed, close to the tentative estimate of $10.8 \text{ J m}^{-2} \text{ s}^{-1}$ by Aagaard and Greisman¹⁴ for the net heat input into the Arctic Ocean attributed to oceanic transport. Estimates for the value of this heat flux vary considerably, with the value chosen here being rather high. This would have produced a conservative result as the ice thickness would have been underestimated. For surface temperatures below 273.1 K it was

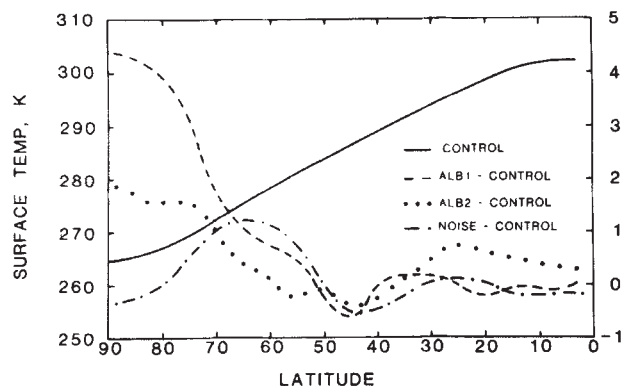


Fig. 2 Zonally averaged surface temperature for the control run, scale on left, and surface temperature differences for the various runs, scale on right. The 'noise' run was a perturbation experiment designed to indicate the natural variability of the model.

assumed that accretion occurred to the underside of the ice floe at the rate

$$DZ = -DT(F_0 - F_i)/k_t$$

The sea ice could, of course, melt depending on the temporal variability of the net heat input at a given gridpoint.

Although this ice model is extremely simple it has reproduced both seasonal and interannual sea ice variations quite capably for both Arctic and Antarctic models (B.G.H. and H. B. Gordon, data not shown).

Two perturbation experiments were performed with the model to determine whether an ice-free 'ocean' centred on the North Pole was viable. The first, ALB1, simply consisted of replacing the existing glacial surface albedos from about 70°N to the pole by oceanic albedos^{15,16} (Fig. 1). Note that even though annual mean surface albedos are displayed in Fig. 1, the difference between the high-latitude values is rather large, hence the importance of albedo changes to the hypotheses mentioned above. In effect this 'ocean' was of zero depth and heat capacity with its surface temperature determined as in the control. Starting from fully developed synoptic distributions from the original model (the control run) the experiment was integrated for 117 model days, with the final 72 days being time averaged and used for analysis.

The second experiment, ALB2, had a slightly larger ocean extending approximately 65°N to the pole. In contrast to ALB1, the thermal capacity of a mixed layer ocean with a uniform depth of 70 m was included in the surface temperature calculation. The initial surface albedos were the same as ALB1. No horizontal currents or transport processes were permitted in this ocean. This run was initiated from the equilibrated fields of ALB1 at day 45, and for the next 26 days the mixed layer ocean was not permitted to cool below 271.2 K. This adjustment period was designed to permit the atmospheric system to conform to the ice free oceanic condition to see whether this condition could be maintained. Subsequently sea ice was allowed to grow with an initial assumed albedo of 0.4, which increased to 0.7 for temperatures below 263 K. The model was integrated for just over 100 days in this mode with the time averaged results for the last 78 days being analysed.

The surface temperature response to the high-latitude albedo changes is shown in Fig. 2. The largest polar warming, 4.5 K in ALB1, was insufficient to raise the surface temperature above the freezing point of seawater. The comparison with the noise model in Fig. 2 indicates that the surface temperature variation in ALB1 was essentially locally confined to the area where the surface albedo was modified. Examination of various general circulation characteristics revealed close similarity in the overall behaviours of ALB1 and the control. For example the zonal mean poleward fluxes of sensible and latent heat owing to large scale eddies varied noticeably from those of the control only polewards of 75°. A hemispherically averaged atmospheric

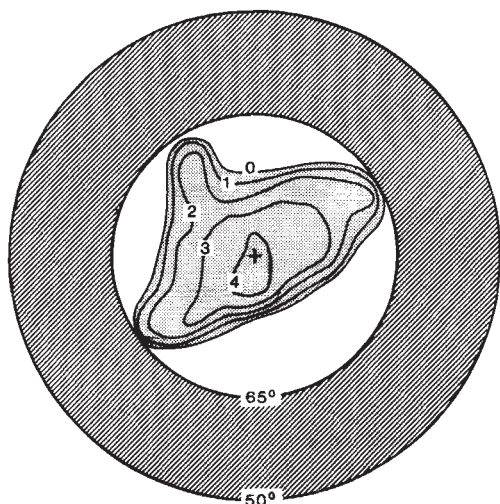


Fig. 3 Sea ice extent and thickness for a nominal Arctic Ocean under annual mean conditions after 92 days model integration. Contour intervals are in metres. The pole is indicated by the + sign. Land is shown hatched, unfrozen ocean is clear and sea ice stippled.

warming of 0.1–0.2 K resulted from the albedo reductions, suggesting that they can make a very modest contribution to a global warming.

In ALB2 the zonal mean surface temperature, Fig. 2, at very high latitudes was only about 1.5 K warmer than the control. ALB2 exhibited rather more variability from the control than ALB1 in its general circulation characteristics, possibly because it had not quite attained equilibrium. For example, the sea ice extent at the end of the experiment, Fig. 3, was still increasing in thickness and, extremely slowly, in spatial extent. The ice distribution was very asymmetric, reflecting the computed sea surface temperatures, but in three locations it extended almost to 65° latitude. Note that this ice cover was grown in less than 100 days under annual mean conditions, hence no difficulty would be expected in achieving such a cover during winter in a seasonally varying model. This ice would then be expected to survive throughout the year, as is observed.

Thus these results indicate that surface albedo associated with continental dispositions at high latitudes is not a decisive factor in determining frozen polar regions. Furthermore the results do not support conjectures that land (of high albedo relative to oceans) is also required at low latitudes for glaciation¹⁷. The low latitude albedo in Fig. 1 is very close to the oceanic value, yet the control, ALB1 and ALB2 runs all attained sub-freezing temperatures at high latitudes. The experiments reinforce the conclusions of Burrett⁵ "that no obviously simple correlation exists between the latitudinal distribution of land and total surface albedo for epochs immediately preceding the initiation of ice-caps . . .".

The above remarks do not imply that high-latitude land plays no role in ice ages. For immense ice sheets such as now exist in Greenland and Antarctica to be produced, it is necessary to have some land, or extremely shallow seas where an ice sheet can bottom, to permit their growth. Thus appropriately located land is a necessary requirement for an ice age to attain its full potential, as noted by Crowell and Frakes⁴, but is not needed for its initiation.

Furthermore, the experiments show that an initial ice free Arctic Ocean would be expected to refreeze, and that an ice extent and thickness similar to that observed would be attained. A stable, alternative ice-free state does not seem viable for overall climatic conditions similar to the present. This conclusion is at variance with the requirements of the well known Ewing–Donn theory of ice ages¹⁸.

The implications of these experiments extend considerably beyond challenging the land/albedo hypothesis. Thus Fig. 3

suggests that the preferred state of the Arctic Ocean is ice covered, unless some substantial climatic fluctuation provides a considerable energy flux convergence into the polar region to provide a warming. This result is eminently reasonable when one considers the geometrical and seasonal factors which minimize the solar input at high latitudes. Clearly during the polar night temperatures well below freezing point are to be expected, the resulting perennial sea ice on the Arctic Ocean would maintain surrounding surface temperatures below or near freezing over much of the year. Most precipitation would occur as snow, and over a sufficient length of time adjacent continental ice sheets would accumulate. The viability of this proposal seems central to the requirements of the Milankovitch theory of ice age perturbations¹⁹. The theory requires that rather small decreases in the geographical distribution of solar radiation cause dramatic ice sheet increases. The presumption is that the associated cooling increases snow accumulation leading to the growth of the existing ice sheets, similar to the above proposals. Thus frozen polar regions seem to be a necessary and sufficient condition (subject to some access to oceans for a source of water vapour) for the creation of ice ages. In particular, there is no necessity to invoke any special mechanism to promote a global cooling in order to initiate an ice age.

If then the past three glacial periods in Earth's history represent the norm the problem is to explain the intervening warm periods, as these should be considered climatic aberrations. This is a problem which has been noticeably neglected, if its existence is, in fact, generally appreciated. It is far easier to devise mechanism for cooling rather than warming the Earth²⁰. Significantly the current proposition simplifies the problem of past climates: now it is only necessary to account for warm periods, previously it was necessary to explain both warm and glacial periods.

One of the critical problems of explaining past warm periods is their equability¹ that is, there was less latitudinal contrast in climate. This feature eliminates many possible dynamical explanations, and the simplest mechanism is that of greatly increased atmospheric CO₂ amounts during such warm periods, as advocated by Budyko²¹. Model calculations^{22,23} support the concept, particularly as regards the equability requirement, as enhanced warmings are produced at high latitudes. Budyko²¹ using a simple zonally averaged heat balance model has computed globally averaged surface temperatures for the past 570 Myr allowing for variations in surface albedo, solar constant and CO₂. He claims remarkably good agreement with observations with temperature increases of up to 10 K for CO₂ enhancements of a factor of 10–12. Additional calculations by Budyko⁸ indicate that such temperature increases, if maintained during winter, would remove the existing Arctic Ocean ice cover over a period of years. Certainly the existence of limestone sediments, coal deposits and so on demonstrates that enormous amounts of CO₂ have effectively been transferred through the atmosphere over geological time. It would be unreasonable to assume that large variations have not also occurred in the atmospheric CO₂ content over this time²⁴.

In essence then it is proposed that ice ages are the natural climatic state of the Earth for atmospheric CO₂ amounts similar to that now existing, and that in accordance with Budyko²¹ warmer periods were associated with much higher levels of CO₂. The elegant simplicity of the proposal has much to commend it. Estimates of past atmospheric CO₂ levels are clearly desirable to evaluate this hypothesis. Other climatic perturbations such as continental drift, volcanic activity and short term solar fluctuations should then be seen as modulators of the basic climatic state, but without the ability to cause transitions to a different state.

Received 21 June; accepted 7 December 1983.

1. Frakes, L. A. *Climates Throughout Geologic Time* (Elsevier, Amsterdam, 1979).
2. Beaty, C. B. *Am. Sci.* **55**, 452–459 (1978).
3. Sellers, A. & Meadows, A. J. *Nature* **254**, 44 (1975).
4. Crowell, J. C. & Frakes, L. A. *Am. J. Sci.* **268**, 193–224 (1970).
5. Burrett, C. F. *Nature* **296**, 54–56 (1982).

6. Fletcher, J. O. *Seventh INQUA Congress, Symp. on Causes of Climatic Change*, Colorado, 1965.
7. Aagaard, K. & Coachman, L. K. *EOS* **56**, 484–486 (1975).
8. Budyko, M. I. *Climate and Life* (Academic, New York, 1974).
9. Fletcher, J. O., Mintz, Y., Arakawa, A. & Fox, T. *World Met. Org., Tech. Note* 1299 (1973).
10. Newson, R. *Nature* **241**, 39–40 (1973).
11. Hunt, B. G. *Mon. Wea. Rev.* **104**, 333–350 (1976).
12. Smagorinsky, J., Manabe, S. & Holloway, J. L. *Mon. Wea. Rev.* **93**, 727–768 (1965).
13. Parkinson, C. L. & Washington, W. M. *J. geophys. Res.* **84**, 311–337 (1979).
14. Aagaard, K. & Greisman, P. J. *geophys. Res.* **80**, 3821–3827 (1975).
15. Posey, J. W. & Clapp, P. F. *Geophys. Inst.* **4**, 33–84 (1964).
16. Hummel, J. R. & Reck, R. A. *J. appl. Met.* **18**, 239–253 (1979).
17. Cogley, J. G. *Nature* **279**, 712–713 (1979).
18. Ewing, M. & Donn, W. L., *Science* **123**, 1061–1066 (1956).
19. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).
20. Barron, E. J., Thompson, S. L. & Schneider, S. H. *Science* **212**, 501–508 (1981).
21. Budyko, M. I. *Met. Hyd. (Soviet)* 5–10 (1981).
22. Manabe, S. & Stouffer, R. J. *Nature* **282**, 491–483 (1979).
23. Parkinson, C. L. & Kellogg, W. W. *Climatic Change* **2**, 149–162 (1979).
24. Kuhn, W. R. & Kasting, J. F. *Nature* **301**, 53–55 (1983).

Is neutron radiation exposure always detrimental to metals (steels)?

K. Linga Murty

Nuclear Engineering Department, North Carolina State University, Raleigh, North Carolina 27650, USA

It is well established that irradiated metals and alloys exhibit increased mechanical strength accompanied by reduced ductility, commonly referred to as irradiation strengthening and embrittlement. We talk about radiation damage and environmental degradation of metals following radiation exposure. Indeed, there have been numerous conferences and symposia held and planned on this subject, which include research work and discussions with the central theme being the damage created in materials by neutron radiation exposure. Radiation embrittlement in metals is believed to be due mainly to (1) changes in flow properties because of the interaction of dislocations with irradiation-produced defects¹, and (2) precipitation of transmutation-produced gases and irradiation-induced segregation at grain boundaries which are potential fracture sites². Contrary to the previously observed radiation embrittlement in metals, I report here that neutron irradiation of mild steel improved both the mechanical strength and ductility at 373 K even at relatively high ($\sim 10^{18}$ neutrons cm^{-2}), fast (>1 MeV) neutron doses. To my knowledge, such improvement has never been observed before.

The material used for the present study was cold-drawn rimmed mild steel whose composition is given in Table 1. Mild steel wires of 0.001 m diameter and 0.0385 m gauge length were annealed *in vacuo* at 973 K to a final grain size of 0.038 mm. Prepared specimens were irradiated to obtain four different fluences: 3.9×10^{16} , 2.8×10^{17} , 2.0×10^{18} and 1.4×10^{19} neutrons cm^{-2} (>1 MeV). The experimental details are given in refs 3 and 4. Special precautions were taken to keep the irradiation temperature constant at 353 K. All of the mechanical tests were performed on a tensile testing machine at a nominal strain rate of $1.36 \times 10^{-4} \text{ s}^{-1}$. Test temperatures were varied from ambient to 573 K; however, we report here the experimental results at room temperature (~ 293 K) and 373 K.

Figure 1 is a compilation of load-elongation curves of mild steel at ambient temperature and illustrates the effect of neutron irradiation resulting in increased tensile strength and pronounced embrittlement, particularly at high fluences. The room temperature ductility of $\sim 40\%$ decreased to $\sim 25\%$ at 2.0×10^{18} neutrons cm^{-2} . At a still higher fluence of 1.4×10^{19} neutrons cm^{-2} a precipitous drop in ductility ($\sim 3\%$) was noted. These results are similar to the earlier observations on radiation embrittlement of metals¹. In addition, the rate of work-hardening decreased while Lüders strain increased with increasing neutron fluence; at the highest fluence, fracture occurred during Lüders propagation itself resulting in pronounced embrittlement.

Similar load-elongation curves of mild steel observed at 373 K are reproduced in Fig. 2 where we note that the unirradiated material exhibited serrations or jerks beyond the Lüders strain in the work-hardening regime, unlike that at room temperature which had a smooth stress-strain curve. This unstable plastic flow is due to the so-called dynamic strain-ageing or blue brittle phenomena occurring in these materials due to the interaction of interstitial impurities such as C and N with dislocations^{3,4}. It is well known that the onset of unstable flow is characterized by a precipitous drop in ductility, as is clear from Fig. 2 where the ductility at 373 K decreased to $\sim 11\%$ from 40% at room temperature. After neutron radiation to 3.9×10^{16} neutrons cm^{-2} , the load-elongation curve exhibited random 'wiggles' in the load in an otherwise smooth curve, but the elongation to fracture increased simultaneously with the stress for plastic flow. As neutron fluence increased, the yield stress (the stress at which Lüders band propagates) increased together with the ductility up to $\sim 2.0 \times 10^{18}$ neutrons cm^{-2} with essentially smooth deformation curves. At the highest fluence of 1.4×10^{19} neutrons cm^{-2} , the yield stress increased with a drop in ductility to $\sim 3\%$,

Fig. 1 Stress-strain curves of mild steel at room temperature depicting the effect of incremental neutron irradiation.

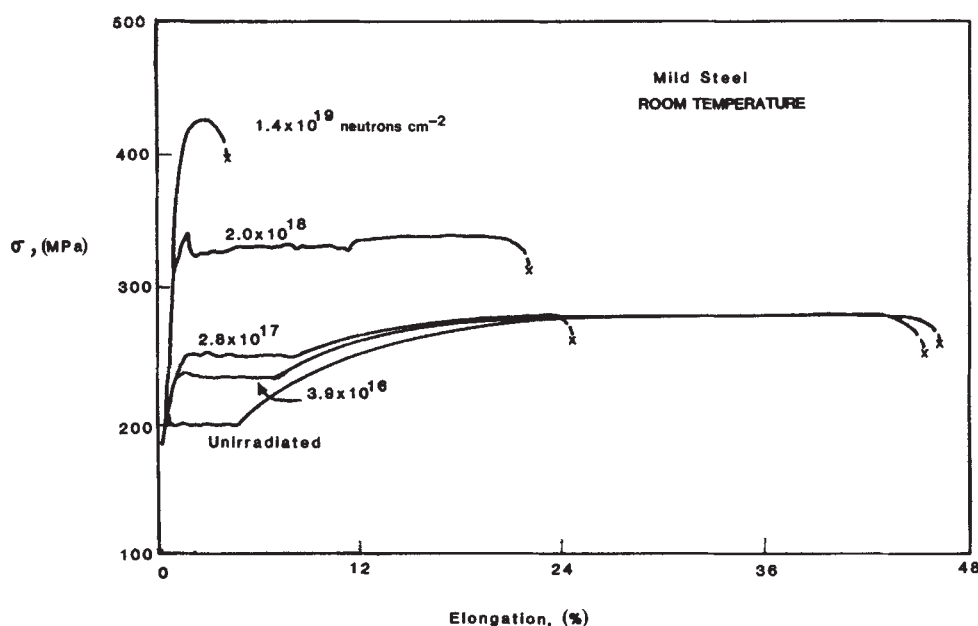
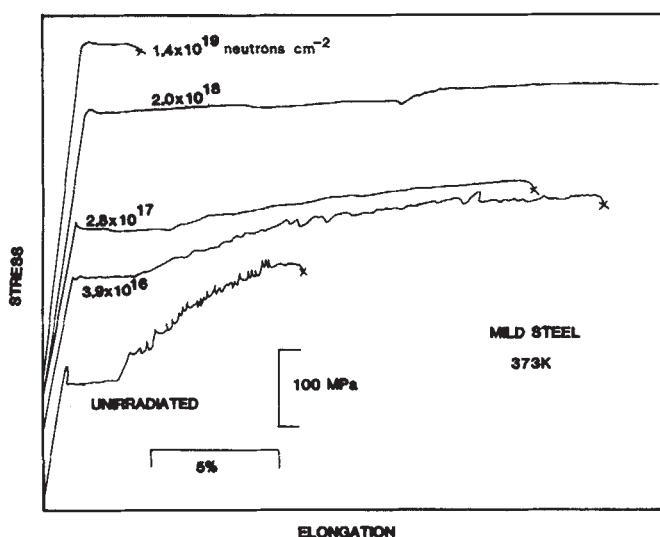


Table 1 Composition of steel wires (wt %)

C	0.050	Mn	0.390
N	0.004	S	0.012
O	0.012	Cr	0.041
Si	<0.001	Cu	0.091
Al	0.002	Sn	0.003
Ni	0.032	Fe	Remainder

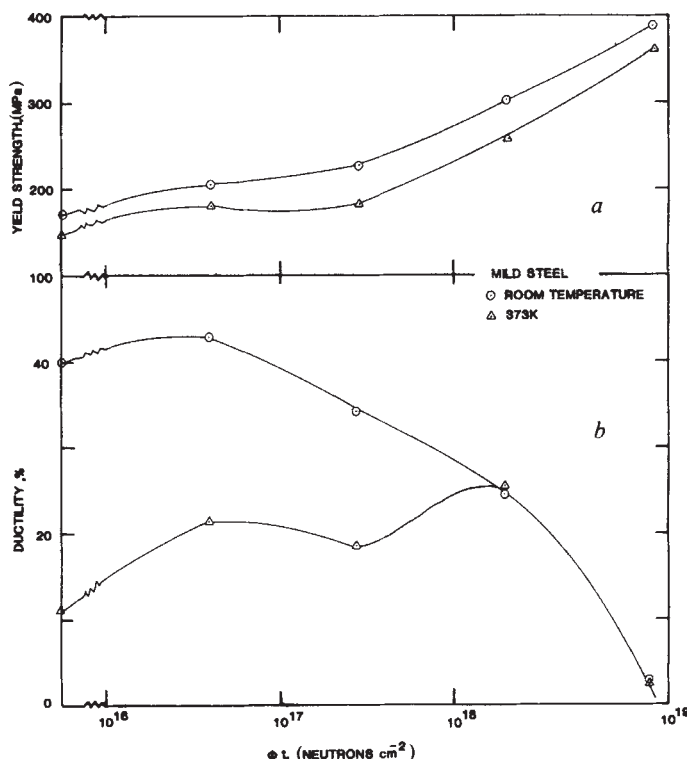
Table 2 Effect of neutron irradiation on strength and ductility at room temperature (RT) and 373 K

Fluence (neutrons cm ⁻²) (>1 MeV)	Yield strength (MPa)		Ductility (%)	
	RT	373 K	RT	373 K
Unirradiated	172.9	147.6	39.8	11.0
3.9×10^{16}	206.4	179.1	42.9	21.3
2.8×10^{17}	227.1	185.0	34.2	18.5
2.0×10^{18}	304.5	259.0	25.1	25.2
1.4×10^{19}	390.2	361.3	3.1	3.0

**Fig. 2** Stress-strain curves of mild steel at 373 K as a function of neutron fluence. Note the jerky load-elongation curve for unirradiated material, with a pronounced drop in ductility. Except at the highest fluence note the improvement in mechanical strength and elongation following radiation exposure.

the same as that at room temperature. The effects of neutron irradiation on yield stress and ductility are shown in Table 2 and are depicted in Fig. 3a and b, respectively. It is clearly demonstrated that at 373 K, the yield stress increased by 75% and the ductility increased by 100%(!) following neutron radiation exposure to 2.0×10^{18} neutrons cm⁻² (>1 MeV). This surprising beneficial effect of neutron irradiation is due to the fact that irradiation-produced defects interact with interstitial impurities (C and N) in steels, resulting in a reduction of the concentration of these impurity atoms available for blue brittleness, thereby increasing the temperature of unstable flow^{3,4}. These synergistic effects of dynamic strain-ageing and neutron irradiation will be discussed in more detail at the Sixth International Conference on Fracture (ICF6) to be held in New Delhi, India in December 1984.

These results have significant impact on our understanding of the radiation embrittlement of nuclear pressure-vessel steels and pressurized thermal shock (PTS) phenomena presently confronting the nuclear (light-water reactor) industry⁵. This type of synergistic effect should also be present in typical nuclear

**Fig. 3** Effect of neutron irradiation fluence on yield stress (a) and ductility (b) of mild steel at room temperature and 373 K.

pressure-vessel steels (for example, A533B and A302B), albeit to a lesser degree due to increased alloying elements. The limited work reported in these steels (unirradiated condition) does demonstrate the existence of dynamic strain ageing⁶. The fact that both the stress and ductility increased in the present study at 373 K following irradiation (to 2.0×10^{18} neutrons cm⁻²) implies that toughness would also increase, contrary to the accepted norm of decrease in toughness after radiation exposure⁵. This is because the rate of deformation used in the present study is quite low compared with that used in obtaining pressure-vessel surveillance data which are usually determined from Charpy impact tests and thus these effects due to strain-ageing would occur at much higher temperatures. In addition, the ranges of temperatures and strain rates where these synergistic effects would be important are functions of alloy composition and heat treatment. Further work on pressure-vessel steels is planned to delineate these differences.

I would like to acknowledge numerous helpful discussions with Professor E. O. Hall, Dr J. Read Holland, and Mr Dirk J. Oh.

Received 16 September; accepted 15 November 1983.

1. Bloom, E. E. in *Radiation Damage in Metals* (eds Petersen, N. L. & Harkness, S. D.) 295-329 (American Society for Metals, Metals Park, Ohio, 1975).
2. Wolfenden, A., Holland, J. R., Lott, R. G. & Spitznagel, J. A. in *Phase Stability During Irradiation* (eds Holland, J. R., Mansur, L. K. & Potter, D. I.) 383-413 (AIME, Warrendale, Pennsylvania, 1981).
3. Murty, K. L. & Hall, E. O. in *Irradiation Effects on the Microstructure and Properties of Metals* 53-71 (ASTM STP 611, Philadelphia, 1976).
4. Murty, K. L. *Mater. Sci. Engng* **59**, 207 (1983).
5. Steele, L. E. in *Status of USA Nuclear Reactor Pressure Vessel Surveillance for Radiation Effects* (ed. Steele, L. E.) 227-272 (ASTM STP 784, Philadelphia, 1983).
6. Ostensson, B. & Westin, R. *The Fracture Toughness of A533B Pressure Vessel Steel at Low Strain-Rates*, Pap. S-573 (Aktiebolaget Atomenergi, Studsvik, Sweden, 1977).

The median valley, a result of magma fracture beneath mid-ocean ridges

Paul M. Davis

Department of Earth and Space Sciences, University of California, Los Angeles, California 90024, USA

The shape of the ocean floor has been successfully modelled using the framework of plate tectonics¹⁻⁴. Lithospheric plates approximately 100 km thick lie above the hotter, less viscous asthenosphere. At the ridges, the lithosphere thins. Plate separation is accompanied by creation of new ocean floor as the asthenosphere flows upwards and cools on approaching the surface. As a result, heat is concentrated beneath the ridge axes. Densities are lower due to thermal expansion, so that for isostatic equilibrium the sea floor sits some 6 km shallower than older sea floor, well off axis, where the temperature anomaly is absent. At fast spreading ridges (velocity $v > 5 \text{ cm yr}^{-1}$) such as the East Pacific Rise, this description fits well. A central ridge crest is flanked on either side by bathymetry which deepens in agreement with thermal models³⁻⁶ at a rate proportional to the square root of age of creation. However at slow spreading ridges, $v < 5 \text{ cm yr}^{-1}$, a major central graben of order 1 km deep replaces the central ridge crest, although the square root of age dependence is observed well away from the axes. Isostatic equilibrium is not achieved⁷, since the sea floor lies 1–2 km below the equilibrium level. In this report I demonstrate that this graben or median valley can result from steady state dyke intrusion into, and brittle failure of, the axial lithosphere.

Figure 1 is modified from a current thermal model^{3,4} of creation of the lithosphere. Solid lithosphere overlies relatively fluid asthenosphere, which accretes onto the lithosphere as it moves away from the ridge axis. Under the ridge crest, molten material from the asthenosphere is introduced episodically into a narrow vertical crack and cools to form dykes. The vertical thickness of the intrusion zone varies inversely with spreading rate. It has been estimated that a ridge spreading at a rate of 1 cm yr^{-1} has an intrusion zone depth of 3.2 km (ref. 4). Geophysical and geological evidence supports this description of the axial lithosphere⁸⁻¹³. The presence of a continuous layer of strength at the ridge axis has been inferred from other thermal as well as petrological models^{8,9}. Rayleigh wave dispersion studies in the Pacific¹⁰ indicate lithosphere thickness at the ridge crest is only a few kilometres. Studies of plate flexure models of isostatic compensation also give thickness values of 2–3 km (refs 11–13).

The field evidence includes mapping of ophiolite complexes, thought to consist of upthrust oceanic crust and upper mantle¹⁴. Layers of sub-vertical sheeted dykes have been found which, observation suggests, formed at depths of 1–3 km (refs 15–16). The dykes are emplaced against each other without intervening country rock and have a thickness of order 1 m. Above the sheeted dykes a layer of pillow lavas is found which resulted from lava extrusion and rapid quenching on the sea floor. Dykes form during an intrusion episode. As the sea floor spreads they are rafted away to form the sheeted layers. We suggest here that the forces of dyke intrusion are responsible for the formation of the median valley.

Surface displacements associated with dyke intrusion or magma-fracture have been measured in the rift zone of Kilauea, Hawaii¹⁷ and north and south of Krafla volcano in northern Iceland¹⁸. Two-dimensional, near-vertical elastic crack models¹⁹ have been used to invert the deformation to estimate the crack geometry and position. Depths to centre, heights and widths, which have values in the range 2–3 km, 1–5 km and 0.5–2 m respectively, are similar to values seen for the sheeted dykes. The surface deformation above a buried dyke consists of two ridges separated by a central graben which lies above the dyke (Fig. 2a). In this study, the important observation is that the

deformation includes inelastic components corresponding to fissures at the ridge crests and inward-dipping normal faults that accentuate the central graben¹⁹. Away from the ridge crests where stresses are lower, the deformation profile is smooth and presumed to arise from elastic strain, allowing the profile to be separated into elastic and anelastic components. The anelastic vertical deformation, which resembles the inferred elastic deformation in form, is estimated to be 1–2 times the amplitude of its elastic counterpart. We can show that steady state superposition of such anelastic profiles at a spreading mid-ocean ridge gives rise to the central graben.

A convenient mathematical description of vertical surface deformation due to magma-fracture into vertical dykes involves approximating the fracture by a dislocation²⁰ with displacement normal to the dislocation surface (Fig. 1). For an infinitely long fracture the vertical displacement, W_e , is given by²¹,

$$W_e(x) = \frac{b}{\pi} \left\{ \frac{d_2^2}{x^2 + d_2^2} - \frac{d_1^2}{x^2 + d_1^2} \right\} \quad (1)$$

where b is the Burger's vector or normal displacement of the crack surfaces, d_1 and d_2 upper and lower crack edges respectively. b is related to the internal overpressure P_0 (that is, pressure above lithostatic), and the modulus of rigidity of the surrounding medium, μ , by²⁰,

$$b \approx \frac{P_0}{\mu} (d_2 - d_1) \quad (2)$$

A representative profile is plotted in Fig. 2a.

We next assume that the anelastic deformation is given by the elastic expression of equation 1. This assumption is made for two reasons. First, as stated earlier the observed anelastic deformation resembles the form of the inferred elastic deformation¹⁹. Second, were we to model the brittle failure of the surrounding medium using the correspondence principle of viscoelasticity²², the viscoelastic form of the resulting deformation would be identical to that of the elastic case. Therefore, after each dyke injection an anelastic deformation $W_e(x)$ is preserved in the ocean floor topography, but rafted away from the ridge crest at the spreading velocity. The effect of superimposed episodes of magma-fracture is then given by,

$$W_a(x) = \int_0^{x/v} \frac{\dot{b}}{\pi} \left\{ \frac{d_2^2}{(x-vt)^2 + d_2^2} - \frac{d_1^2}{(x-vt)^2 + d_1^2} \right\} dt \quad (3)$$

$$= \frac{\dot{b}}{\pi v} \left\{ d_2 \tan^{-1} \left(\frac{x}{d_2} \right) - d_1 \tan^{-1} \left(\frac{x}{d_1} \right) \right\} \quad (4)$$

where $\dot{b}$ is the steady state dislocation rate, or the excess injection rate into the dyke zone over that required to fill the gap caused by spreading. This results in a graben (Fig. 2b) with width many times that of an individual dyke. Since the graben is depressed below isostatic equilibrium for the ridge problem, we take $W_a(x=\infty)$ as the level at which isostasy is achieved.

The final step in modelling the ridge crest structure is to superimpose the magma-fracture model onto the topography predicted by lithosphere cooling models^{5,6} that is

$$W_c(x) = -e - f\sqrt{x/v} \quad (5)$$

where e has been found to have values of $2,500 \pm 250 \text{ m}$ and f ranges from 0.33 to $0.52 \text{ km Myr}^{-1/2}$ (ref. 5). This range implies regional variation in the plate parameters: thickness, density, thermal diffusivity or base temperature. Thus we have no *a priori* expectations for these values for a given situation, other than they are expected to fall in the range found for the fastest spreading ridges where the median valley, and by implication, magma-fracture effects are absent.

The cooling model topography is plotted in Fig. 2c. Near the ridge crest, the cusp at $x=0$ due to the $\sqrt{x/v}$ dependence has been smoothed over the range $|x| < 2 \text{ km}$ using a quadratic in x , to give a curve close to that predicted by the intrusion zone thermal models^{3,4}.

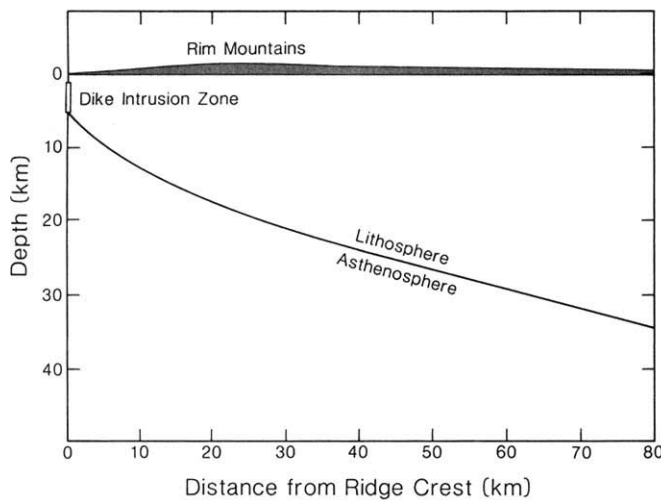


Fig. 1 Shape of the upper and lower surfaces of the lithosphere for a slow spreading ridge. The dyke injection zone transfers upward forces of buoyancy to lateral forces of intrusion which results in ridges flanking a central graben which is lower than the isostatic compensation level. Each dyke is modelled as an infinitely long vertical crack of rectangular (dislocation) cross-section beneath the surface of a semi-infinite medium. Depths of the upper and lower crack edges are d_1 and d_2 respectively.

The composite model is then given by

$$W(x) = -e - f\sqrt{x/v} + \frac{b}{\pi v} \cdot \left\{ d_2 \tan^{-1} \left(\frac{x}{d_2} \right) - d_1 \tan^{-1} \left(\frac{x}{d_1} \right) \right\} - \frac{b(d_2 - d_1)}{2v} \quad (6)$$

where the last term represents the adjustment of the fracture model to isostatic equilibrium. $W(x)$ is plotted in Fig. 2d fitted to an averaged bathymetry profile from the Mid-Atlantic Ridge²³. Profiles between 14°10' and 14°40' N were stacked and averaged and the resulting profile symmetrized²³. The stacking and averaging reduced the roughness in the topography, a predominant feature of slow spreading ridges. Values for the parameters used in the fit are $d_1 = 1$ km, $d_2 = 5$ km, $b = 0.80$ cm yr⁻¹, $v = 1$ cm yr⁻¹, $e = 2,250$ m, $f = 0.52$ km (Myr)^{-1/2}. The values of e and f for the cooling model fall to one side of but within the range found for ridges without valleys⁵.

Similar values have been observed at one of the fastest spreading axes, that of the southern East Pacific Rise ($v = 7$ cm yr⁻¹) (ref. 5). A less steep cooling model may be applicable (for example that of Parsons and Sclater⁶) if the magma fracture effects are more confined to the axis than predicted by equation 4. Such strain concentration would occur if the newly formed lithosphere is significantly weaker at the ridge axis than away from it. Since the strength variation is unknown, this more complex problem is unsolvable at present.

For the dislocation model, the depth, $(d_1 + d_2)/2$, is the most constrained parameter. An inverse trade-off between height, $d_2 - d_1$, and injection rate, b , occurs which gives minor changes to the fit. The chosen value for height spans oceanic layer 2B but is larger than the measured heights of the sheeted dykes of the ophiolite complexes. The injection rate is a measure of the anelastic failure of the medium surrounding the dyke. Since the central graben is depressed some 1–2 km below the level of the isostatic compensation, magma will inject into the dyke at an over pressure of 200–400 bar, causing a distension greater than that required to replenish the material lost due to plate separation. Assuming conservation of lateral flow, the increased injection at the dyke zone balances the increased cross-section to the flow at the rim mountains. Therefore, from a steady state

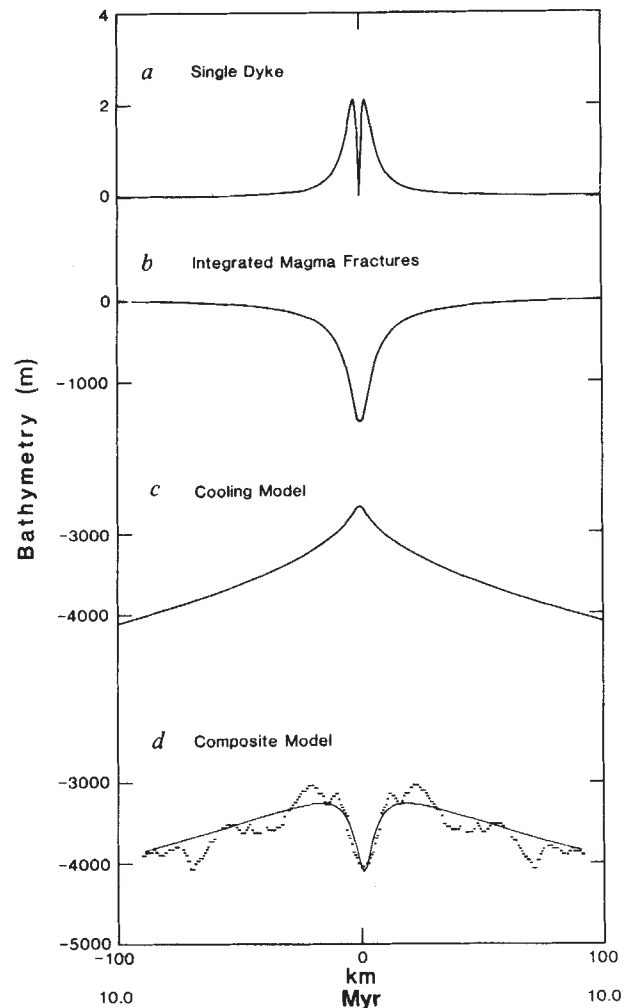


Fig. 2 *a*, Elastic or viscoelastic deformation due to injection of a dyke beneath the surface of a half-space. *b*, Superposition of the deformation *a* due to steady state dyke intrusion into a spreading lithosphere. The large graben which results is taken to lie below the level of isostatic equilibrium for the mid-ocean ridge case. *c*, Isostatically compensated topography expected from plate cooling models in the absence of dyking or constricted flow of the asthenosphere at the ridge axis. *d*, Superposition of curves *b* and *c* and comparison to digitized (dots) averaged bathymetry (after ref. 23) from the Mid-Atlantic Ridge.

point of view, the excess injection and dyke solidification deflects the upward flow at the ridge crest, into the rim mountains. If the medium remained elastic, equation 4 describes the resulting surface deformation which must sustain strains of order 10^{-1} . Elasticity will break down well before this, as seen in the Hawaiian and Icelandic examples.

Dyke injection into rift zones and associated volcanism is episodic. For example, rifting of the plate boundary in northern Iceland occurs every 100–150 years¹⁸. In this model, during the period between injection episodes, the stresses will relax in two ways: through anelastic failure and flow as seen on Krafla and Kilauea, and by relative plate motion. If the former is absent, injection episodes will keep stresses at a steady level if the plates separate by a distance equal to the dyke width. Then there will be no excess injection or rim mountain formation. The elastic deformation associated with the initial injection will be unstrained since the dislocation is accommodated through plate motion. However, if both mechanisms occur and the former is 1–2 times the amplitude of the latter, stresses will be relaxed over a much smaller interval, allowing the next episode to occur much sooner. The excess injection rate, which is 40% of the spreading rate, can be explained by this process. The Hawaiian

and Icelandic observations support the argument for anelastic failure of this magnitude.

Although one must treat with caution a model involving a 5-parameter fit, we would counter that the parameters are all within the range observed for these processes elsewhere.

The depth to the centre of the dyke zone is governed by the width (approximately one tenth) of the median valley. For faster spreading ridges, this width decreases. We have fit the model to bathymetric profiles⁷ from the North ($v = 1.20 \text{ cm yr}^{-1}$) and South ($v = 1.5 \text{ cm yr}^{-1}$) Atlantic and the Sheba ($v = 1.4 \text{ cm yr}^{-1}$) and Gorda ($v = 2.9 \text{ cm yr}^{-1}$) ridges which gave depths of 3.1, 3.1, 2.6, 1.6 km respectively. When spreading is so fast that no valley is formed, dyke intrusion is not then a determining factor. This leads us to suggest that the difference between fast and slow spreading topographies depends on the ratio of intrusion to extrusion. For slow spreading, the cooler intrusion zone traps a high proportion of injected magma by cooling and solidification, possibly by hydrothermal cooling, before lava can erupt onto the sea floor. (Some eruptions must occur because of the presence of the pillow basalts layer.) In the hotter more mobile lithospheres at fast spreading ridges, extrusion occurs more readily and buoyant rise to isostatic equilibrium is achieved right at the ridge crest.

The model described here incorporates many features of earlier models of the median valley. These include constriction of the upwelling asthenosphere due to viscous forces^{24,25}, necking of the upper brittle/elastic zone⁹, asthenospheric flow due to surface unloading as the plates part²³, and imbalance between dyke injection and lateral strain^{26,27}. This model differs in that constriction occurs in the upper 5 km of the lithosphere where the flow freezes solid. Below this, where temperatures are near or even above the solidus^{3,4}, forces due to viscous flow are assumed to be negligible. Necking in the region above a dyke results from the forces on its surfaces. Owing to the large area subtended horizontally, the major forces are laterally outward. Immediately above the dyke there is a concentrated horizontal tension which, unless the dyke solidifies, enables it to extend upwards by crack propagation. Above the dyke near the surface, this zone of tension bifurcates into two mounds separated by the central graben.

Our model does not fit the high frequency fluctuations in the bathymetry: notably the rim mountains are higher than the smooth theoretical curve. However there are other deviations of comparable magnitude. It is therefore not compelling at this stage to add another ingredient such as plate flexure⁸. The misfit at the rim mountains was even more pronounced for the fits to the other ridge profiles, but again there were other equally large deviations. The fact that the misfit lessened for the averaged profile suggests that the deviations are incoherent on the different profiles and that the description of the first order features of the ridge requires using averaged depth.

The free air gravity anomaly for a ridge in isostatic equilibrium is positive over the ridge crest and decays with distance in a manner similar to the fall-off in the height of the ridge topography. Values of $\sim 60 \text{ mgal}$ have been estimated above the ridge relative to that measured some 300 km away⁴. Superposition of an uncompensated 1,500 m deep axial graben, as in Fig. 2b, would depress the free air anomaly to -40 mgal at the axis. This qualitative description agrees with the observational evidence^{7,23}. Quantitative comparison will require evaluation of the isotherms for the thermal model and integration of the gravity equation, but this is outside the scope of this report.

The composite model proposed here explains the graben, or median valley, and rim mountains found at slow-spreading oceanic ridges. The model superimposes anelastic deformation due to dyke magma-fracturing onto the isostatic topography predicted from plate cooling theory. The success of early thermal models^{1,2} resulted from describing ridge features by understanding near-surface processes, so that deeper motions in the mantle could eventually be unravelled. In the same vein, this model finds the median valley to be a consequence of very-near-surface

effects in the cooled upper 5 km of the lithosphere, rather than the result of flow in deep conduits²⁸ or magma chambers at greater depth²⁹.

Received 22 August; accepted 29 November 1983.

- McKenzie, D. P., *J. geophys. Res.* **72**, 6261–6273 (1967).
- McKenzie, D. P. & Sclater, J. G. *Bull. Volcanol.* **33**, 101–117 (1969).
- Parker, R. L. & Oldenburg, D. W. *Nature* **242**, 137–139 (1973).
- Oldenburg, D. W. *Geophys. J. R. astr. Soc.* **43**, 425–451 (1975).
- Davis, E. E. & Lister, C. R. B. *Earth planet. Sci. Lett.* **21**, 405–413 (1974).
- Parsons, B. & Sclater, J. G. *J. geophys. Res.* **82**, 803–827 (1977).
- Watts, A. B. in *Continental and Oceanic Rifts*, Geodynamic Series No. 8 (ed. Palmason, A. G.) 99–104 (AGU, Washington DC, 1982).
- Turcotte, D. L. *J. geophys. Res.* **79**, 2573–2577 (1974).
- Tapponnier, P. & Francheteau, J. *J. geophys. Res.* **83**, 3955–3970 (1978).
- Leeds, A. R. thesis, Univ. Calif., Los Angeles (1974).
- McKenzie, D. P. & Bowin, C. *J. geophys. Res.* **81**, 1903–1915 (1976).
- Cochran, J. R. *J. geophys. Res.* **84**, 4713–4729 (1979).
- McNutt, M. *J. geophys. Res.* **84**, 7589–7598 (1979).
- Gass, I. G. *Nature* **220**, 39–42 (1968).
- Williams, H. *Geol. Surv. Can. Pap.* 72–34, 1–7 (1973).
- Salisbury, M. H. & Christensen, N. I. *J. geophys. Res.* **83**, 805–817 (1978).
- Duffield, W. A., Jackson, D. B. & Swanson, D. A. in *Proc. Sym. Andean and Antarctic Volcanology Problems 577–587* (Giannini and Figli, Napoli, 1976).
- Björnsson, A., Johnsen, G., Sigurdsson, S., Thorbergsson, G. & Tryggvason, E. *J. geophys. Res.* **84**, 3029–3037 (1979).
- Pollard, D. D., Delaney, P. T., Duffield, W. A., Endo, E. & Okamura, A. T. *Tectonophysics* **94**, 541–584 (1983).
- Davis, P. M. *J. geophys. Res.* **88**, 5826–5834 (1983).
- Maruyama, T. *Bull. Earthquake Res. Inst.* **42**, 289–368 (1964).
- Flügge, W. *Viscoelasticity* (Springer Verlag, New York 1975).
- Collette, B. J., Verhoef, J. & Mulder, A. F. J. *J. Geophys.* **47**, 91–98 (1980).
- Sleep, N. H. *J. geophys. Res.* **74**, 542–549 (1969).
- Sleep, N. H. & Biehler, S. *J. geophys. Res.* **75**, 2748–2752 (1970).
- Deffreyes, K. S. in *Megatectonics of Continents and Oceans* (eds Johnson, H. & Smith, B. L., 1922) (Rutgers University Press, New Brunswick, New Jersey 1970).
- Anderson, R. N. & Noltimer, H. C. *Geophys. J. R. astr. Soc.* **34**, 137–147 (1973).
- Lachenbruch, A. H. *J. geophys. Res.* **81**, 1883–1902 (1976).
- Sleep, N. H. & Rosendahl, X. X. *J. geophys. Res.* **84**, 6831–6839 (1979).

Moho beneath the North Sea compared on normal incidence and wide-angle seismic records

Penny Barton*, Drummond Matthews*,
Jeremy Hall† & Michael Warner*

* Bullard Laboratories, Madingley Road,
Cambridge CB3 0EZ, UK

† Department of Geology, University of Glasgow,
Glasgow G12 8QQ, UK

The Mohorovicic discontinuity (the Moho) is properly defined as the step in the velocity of horizontally travelling P waves from about 7 km s^{-1} to about 8.1 km s^{-1} . Its detailed structure varies from place to place, but it is universally interpreted as the transition from crust to mantle. Normal incidence reflections originating at about Moho depth have been reported sporadically during the past 20 yr. During the past 10 yr continuous seismic reflection profiles have been obtained to this depth both on the continents and in the deep sea and, at the same time, striking improvements have been made in the methods available for interpretation of 'refraction' experiments. Here we report a comparison between the Moho determined by a long range refraction experiment across the North Sea and the results of a short normal incidence reflection profile shot along the same line (Fig. 1). We believe this to be the first time that such a direct comparison has been made.

The refraction experiment^{1,2} was completed in 1981 as part of a programme to check the application of the stretching hypothesis³ to the formation of the North Sea basin. Seismograms were interpreted by amplitude modelling, allowing for a velocity structure that varies both vertically and horizontally. The modelling program⁴ ascribes a constant velocity within a 1-km square box; the North Sea model has 53×70 such boxes. The velocities of the central part of the final model are summarized in Fig. 2a. The model is well constrained in this

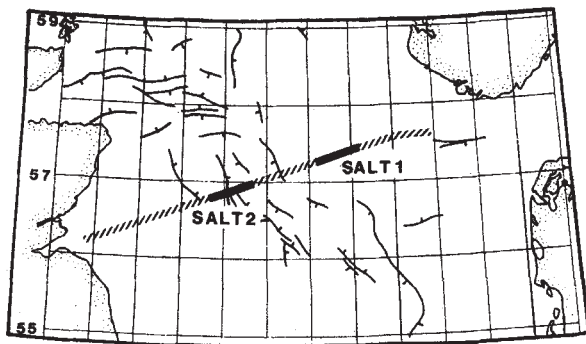


Fig. 1 Map of North Sea showing graben system and position of seismic profiles. Slashed line, refraction profile; solid lines, deep normal incidence reflection profiles.

central region. These velocities were used to convert the model depth-to-Moho into the equivalent normal incidence two-way travel time section. The uncertainty from the modelling of this estimate of the Moho reflection time is about 0.2 s. If the basement were strongly anisotropic (say 5%, with a fast velocity in the horizontal direction) the normal incidence Moho reflection time might be overestimated from the refraction data, but by no more than about 0.2 s.

The seismic reflection profile was acquired for the British Institutions Reflection Profiling Syndicate (BIRPS) in February 1983. It was shot and processed by Western Geophysical as one of two such short profiles (Fig. 1) planned to test whether it is possible to obtain deep reflections despite the thick sediments and salt in this part of the North Sea. A line drawing of the data, based only on reflector strength and continuity, is shown in Fig. 2b. At the west end of the line, west of the marginal fault of the Central Graben, layering is apparent in the lower part of the crust. This layering appears similar to that already reported from the north-west of Britain⁵, from the USA⁶, from Australia⁷ and from Germany (R. Meissner, personal communication), although it occupies less of the lower crust than it does on much of the WINCH line from north-west Britain. In Fig. 2c the Moho reflection times calculated from the refraction model are shown superimposed on the reflection data. The width of the Moho trace is some measure of the uncertainties. It can be seen that the refraction Moho lies along the base of the correlatable reflections from the lower crust. At this place in the North Sea there is no discrete reflector to mark the top of the upper mantle.

These observations are compatible with predictions⁸ of a lower crust that includes laminae of ultramafic rock just above the Moho which are expected to give rise to significant reflections. The laminae are too thin to transmit P waves horizontally at mantle velocities; Moho refracted phases travel through a relatively homogeneous ultramafic layer below the laminae.

We conclude that in this small region of the North Sea, the Moho lies at the base of the layering in the lower crust; no single reflector marks its position. With the increasing international interest in the deep structure of the continents it is important that similar in-line reflection and refraction experiments are undertaken, so that two-dimensional (and eventually three-dimensional) models of seismic velocity and acoustic impedance (density $\times$ velocity) may be compared and combined to give a more complete picture of the deep crust.

The SALT data may be obtained at cost of reproduction from Marine Geophysics Unit, Institute of Geological Sciences, Murchison House, West Mains Road, Edinburgh EH9 3LA, UK from one year after acquisition. The work was supported by

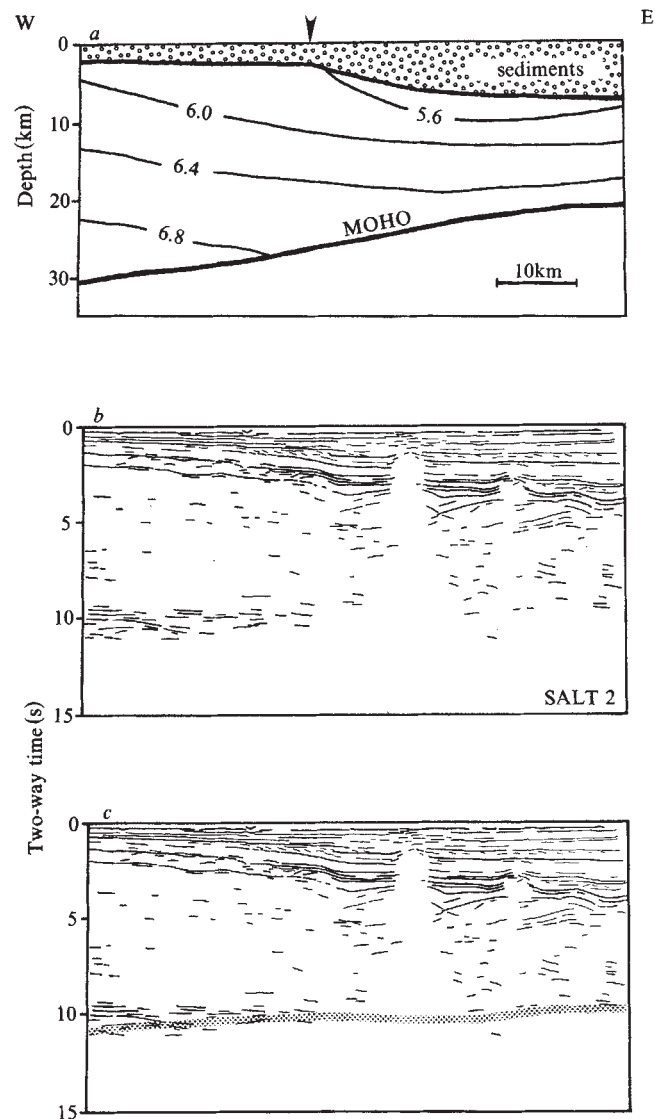


Fig. 2 a, Velocity-depth-distance model of part of refraction profile coincident with SALT2 deep reflection profile. Fine lines show contours in seismic velocity; they do not represent layering. Arrow marks western boundary fault of the Central Graben. b, Line drawing of SALT2 data based on reflector strength and continuity, to same horizontal scale as a. Note band of strong reflectors at about 10 s two-way time on the western part of the line. c, Two-way time normal incidence profile to Moho (stippled line) determined from refraction model of a, superimposed on b. Width of stippled trace is a measure of probable error in the method.

NERC (grant GR3/3932 and BIRPS). J.H. thanks Shell UK for support enabling him to work with BIRPS. Cambridge Earth Sciences Contribution No. 439.

Received 13 October; accepted 16 November 1983.

1. Wood, R., & Barton, P. *Nature* **302**, 134-136 (1983).
2. Barton, P., & Wood, R. *Geophys. J.R. astr. Soc.* (submitted).
3. McKenzie, D. P. *Earth planet. Sci. Lett.* **40**, 25-32 (1978).
4. Cassell, B. R. *Geophys. J.R. astr. Soc.* **69**, 339-354 (1982).
5. Brewer, J. *et al. Nature* **305**, 206-210 (1983).
6. Oliver, J., Cook, F. & Brown, L. *J. geophys. Res.* **88**, B4, 3329-3347 (1983).
7. Mathur, S. P. *First Break* **1**(7), 9-16 (1983).
8. Hale, L. D. & Thompson, G. A. *J. geophys. Res.* **87**, B6, 4625-4635 (1982).

Significance of Mn(IV) oxide in abiotic formation of organic nitrogen complexes in natural environments

H. Shindo* & P. M. Huang

Saskatchewan Institute of Pedology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada S7N 0W0

* Permanent address: Faculty of Agriculture, Yamaguchi University, Yamaguchi, Japan 753

The high resistance of organic nitrogen complexes in soil to microbial attack is of significance to crop production, N balance of soils, geochemical and pedological processes, and environmental quality¹⁻³. Several theories have been given to account for this phenomenon. An important theory indicates the stabilization of nitrogen through the formation of nitrogenous polymers during the reaction of nitrogenous substances with polyphenols¹⁻³. It is well known that the formation of nitrogenous polymers from nitrogenous substances and polyphenols is accelerated in the presence of phenol oxidase enzymes¹⁻⁸ and oxidants such as Ag₂O (ref. 9). However, the promoting effects of soil inorganic components on the formation of organic N complexes have received little attention. Mn(IV) oxides are common in soil¹⁰. Furthermore, the effect of Mn(IV) oxides on the oxidative polymerization of hydroquinone is much greater than that of various other inorganic components such as short-range ordered Fe(III), Al and Si oxides¹¹, primary minerals¹² and clay minerals¹³ which are commonly present in soil environments. We report here that Mn(IV) oxide (birnessite), which is common in natural environments, greatly promotes the abiotic formation of nitrogenous polymers in hydroquinone-glycine systems in the common pH range (4-8) of soils. The findings indicate that Mn(IV) oxides merit close attention in the abiotic formation of organic N complexes from nitrogenous substances and polyphenols and the subsequent turnover of N in soil and the associated environments.

Birnessite (δMnO_2) was prepared and used in this study, since it is one of the most common Mn oxides present in soil¹⁰. Glycine, an amino acid, and hydroquinone, a diphenol, of certified reagent grade were used, because these compounds and their structural units are widely distributed in nature^{1-3,14}. The reaction of glycine with hydroquinone was carried out at the molar ratio of 5:1 (Table 1). Haider *et al.*⁴ used the same ratio to study the reaction of amino acids with diphenols in the presence of phenol oxidase enzymes.

Table 1 shows the concentrations of NH₃-N (ammonia-N released by deamination) and polymer-N (N incorporated into the polymers) formed in the supernatant in the hydroquinone-glycine system both in the absence and presence of Mn(IV) oxide at different pH values at the end of the reaction period

of 24 h. The addition of Mn oxide greatly accelerated the formation of NH₃-N and polymer-N in the hydroquinone-glycine systems at any given pH (Table 1). The formation of NH₃-N and polymer-N in the Mn oxide-hydroquinone-glycine system were influenced by the pH value of the systems. In the initial pH range of 4.0 to 7.3 and in the absence of MnO₂, the concentration of NH₃-N was higher than that of polymer-N. At the initial pH of 6.0 and 7.3 and in the presence of MnO₂, the reverse was true. In the presence of MnO₂, the concentration of polymer-N was substantially higher at the initial pH of 6.0 and 7.3 than at pH 4.0, but there was no significant difference in NH₃-N.

In this study, 700 parts per million (p.p.m.) of N as glycine was added to each system. For example, at the initial pH of 7.3 in the Mn oxide-hydroquinone-glycine system, the yields of NH₃-N and polymer-N were observed to be highest; the concentrations of NH₃-N and polymer-N were 34.1 and 41.2 p.p.m., respectively (Table 1). These data reveal that 4.9 and 5.9% of the original glycine-N were transformed to NH₃-N and polymer-N, respectively, in only 24 h.

According to existing literature, amino acids combine easily with quinones^{1-4,15} which are produced by the oxidation of polyphenols, and the reaction products may polymerize to form nitrogenous polymers¹⁻⁴. During this reaction, part of the amino acids which combined with the quinone C=O group may be deaminated¹⁻⁴. These mechanisms were proposed in the system involving phenol oxidase enzymes. In the present study, the formation of NH₃-N and polymer-N in the hydroquinone-glycine system was greatly accelerated by the presence of Mn oxide (Table 1). The presence of MnO₂ increased the yields of NH₃-N and polymer-N from 2.9 to 6.9 times and from 4.2 to 20.1 times, respectively. The processes involved are proposed as follows: (1) the Mn oxide acts as a Lewis acid by accepting electrons from hydroquinone which is thus oxidized and subsequently polymerized¹¹. (2) The reaction products of glycine with quinones polymerize to form nitrogenous polymers, thereby incorporating glycine into the polymers during the oxidative polymerization of hydroquinone. (3) A partial deamination of glycine which was combined with quinone C=O occurs during process (2). Therefore, the data (Table 1) indicate that MnO₂ has a similar role to phenol oxidase enzymes¹⁻⁴, which act as electron acceptors in promoting the formation of nitrogenous polymers and the deamination of amino acids in diphenol-amino acid systems.

The concentration of polymer-N in the Mn oxide-hydroquinone-glycine system was much higher at the initial pH of 6.0 and 7.3 than at pH 4.0, but this trend was not observed in the absence of Mn oxide (Table 1). In the Mn oxide-hydroquinone-glycine systems to which were added the solutions of initial pH values of 4.0, 6.0 and 7.3, the final pH values after 24 h were 4.1, 7.6 and 7.9 (Table 1). The final pH values in the control systems without Mn oxide were 4.1, 6.1 and 6.9,

Table 1 Formation of NH₃-N and polymer-N (in p.p.m.) in the hydroquinone-glycine system both in the absence and presence of Mn oxide at different pH values at the end of 24 h

	Initial pH of 4.0				Initial pH of 6.0				Initial pH of 7.3			
	Final pH	NH ₃ -N	Polymer-N	Sum-N	Final pH	NH ₃ -N	Polymer-N	Sum-N	Final pH	NH ₃ -N	Polymer-N	Sum-N
Hydroquinone												
Glycine	4.1	4.4	3.0	7.4	6.1	6.0	2.0	8.0	6.9	11.7	3.4	15.1
Mn oxide												
Hydroquinone												
Glycine	4.1	30.5	12.6	43.1	7.6	30.7	40.1	70.3	7.9	34.1	41.2	75.3

20 mg of Mn(IV) oxide (birnessite)¹⁰ was suspended in 20 ml of 0.1 M sodium acetate solution at different pH values (adjusted with dilute acetic acid) containing 200 μmol of hydroquinone and 1,000 μmol of glycine. The flasks (125 ml) containing the suspensions were placed in an oscillating water bath at 25 °C. The control solution (hydroquinone and glycine) without Mn oxide was also placed in the same conditions. At the end of the reaction period of 24 h, the reaction solutions were centrifuged at 38,000g for 20 min. The supernatant obtained was used in the determination of NH₃-N formed. The supernatant was also dialysed against water until the dialysed water was colourless. The dialysis tube (~3,500 molecular weight cutoff, from Spectra Medical Industries Inc., Los Angeles) was used. Dialysis eliminated unreacted glycine and hydroquinone, lower molecular weight polymers and acetate. The materials retained in the dialysis tube was used for the determination of polymer-N. NH₃-N is the ammonia-N formed, which was determined by steam distillation using MgO (ref. 18). Polymer-N is nitrogen that is present in the fraction with a molecular weight of ~3,500 and higher, which was determined by Kjeldahl method¹⁸. Sum-N = NH₃-N + Polymer-N.

respectively. Therefore, the much higher concentration of polymer-N in the Mn oxide-hydroquinone-glycine system in near neutral conditions than in acidic conditions can be attributed to: (1) the formation of nitrogenous polymers is promoted by the Mn oxide at higher pH conditions, and (2) the formation of nitrogenous polymers which is driven by the increase of pH as a result of the reduction of Mn oxide and also as a result of the formation of $\text{NH}_3\text{-N}$ through the deamination. The reduction of MnO_2 consumes H^+ in larger amounts than that released by the oxidation of diphenols, $\text{C}_6\text{H}_4(\text{OH})_2$ (refs 11, 16).

Shindo and Huang reported that the oxidative polymerization of hydroquinone as influenced by Mn oxide occurred abiotically¹¹. In the present study, in order to determine whether microorganisms influence the formation of $\text{NH}_3\text{-N}$ and polymer-N, the darkening supernatant obtained from the Mn oxide-hydroquinone-glycine system at the initial pH of 6.0 at the end of 24 h was inoculated on agar plates as reported elsewhere¹¹, and incubated at 27 °C. The agar plates did not show any growth of microorganisms even after 6 days. Phenolic compounds are known to be toxic to microorganisms at concentrations more than 100 p.p.m. (ref. 17). The data thus indicate that the formation of $\text{NH}_3\text{-N}$ and polymer-N observed in this study occurs abiotically.

The present results reveal that Mn oxide greatly promotes the abiotic formation of $\text{NH}_3\text{-N}$ and polymer-N in the hydroquinone-glycine system in the pH range common in soils, and indicate that this promotion deserves close attention, as do subsequent N transformations in soil and the associated environments.

This study was supported by grants from the Ministry of Education of Japan and from the NSERC of Canada. We thank Dr J. J. Germida of our institute for help with the microbiological aspects of the experiment.

Received 15 August; accepted 17 November 1983.

1. Bremner, J. M. in *Soil Biochemistry* Vol. 1 (eds McLaren, A. D. & Peterson, G. H.) 19-66 (Dekker, New York, 1967).
2. Parsons, J. W. & Tinsley, J. in *Soil Components* Vol. 1 (ed. Gieseking, J. E.) 263-304 (Springer, New York, 1975).
3. Stevenson, F. J. in *Nitrogen in Agricultural Soils* (ed. Stevenson, F. J.) 1-42, 67-122 (American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, 1982).
4. Haider, K., Frederick, L. R. & Flaig, W. *Pl. Soil* **22**, 49-64 (1965).
5. Flaig, W., Beutelspacher, H. & Rietz, E. in *Soil Components* Vol. 1 (ed. Gieseking, J. E.) 1-211 (Springer, New York, 1975).
6. Verma, L., Martin, J. P. & Haider, K. *Soil Sci. Soc. Am. J.* **39**, 279-284 (1975).
7. Haider, K., Martin, J. P. & Filip, Z. in *Soil Biochemistry* Vol. 4 (eds Paul, E. A. & McLaren, A. D.) 195-244 (Dekker, New York, 1975).
8. Martin, J. P. & Haider, K. *Soil Sci. Soc. Am. J.* **44**, 983-988 (1980).
9. Ladd, J. N. & Butler, J. H. A. in *Rep. FAO/IAEA tech. Meet.* 143-150 (Pergamon, Oxford, 1966).
10. McKenzie, R. M. *Miner. Mag.* **38**, 493-502 (1971).
11. Shindo, H. & Huang, P. M. *Nature* **298**, 363-365 (1982).
12. Shindo, H. & Huang, P. M. *Soil Sci.* (submitted).
13. Kumada, K. *Chemistry of Soil Organic Matter* 2nd edn (in Japanese, Japan Scientific Society Press, Tokyo, 1981).
14. Harborne, J. B. & Simmonds, N. W. in *Biochemistry of Phenolic Compounds* (ed. Harborne, J. B.) 77-127 (Academic, London, 1964).
15. Mason, H. S. *Nature* **175**, 771-772 (1955).
16. Weast, R. C. (ed.) *CRC Handbook of Chemistry and Physics* 59th edn (CRC Press, Boca Raton, Florida, 1978).
17. Kuwahara, M., Shindo, N. & Munakata, K. *J. agric. Chem. Soc. Japan* (in Japanese) **44**, 169-174 (1970).
18. Bremner, J. M. in *Methods of Soil Analysis* Pt 2 (eds Black, C. A., Evans, D. D., White, J. L., Ensminger, L. E. & Clark, R. E.) 1149-1237 (American Society of Agronomy, Madison, 1965).

Influence of citric acid on the natural formation of imogolite

K. Inoue* & P. M. Huang

Saskatchewan Institute of Pedology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada S7N 0W0

Imogolite, a hydrous aluminosilicate with paracrystalline cylindrical structures¹, has been extensively isolated from the B horizon of ando soils²⁻⁴ and podzols⁵⁻⁹. Imogolite imparts significant effects on ion exchange properties, acidity, stability and biodegradability of organic matter, and physical and engineering properties of soils^{3,10}. Its formation and stability are therefore of great interest in soil and environmental sciences. Imogolite was synthesized from a solution containing hydroxy-Al ions and orthosilicic acid at pH < 5 (refs 11-13). The interaction of hydroxy-Al ions with orthosilicic acid may be influenced by low molecular weight organic acids which are constantly introduced to soils through natural vegetation and farming. However, relatively few data are available on the impeding effect of low molecular weight organic acids on the formation of imogolite¹⁴ and on the properties of the reaction products. We report here that citric acid (molar ratio of citric acid to Al at ≤ 0.1) greatly perturbs the interaction of hydroxy-Al ions with orthosilicic acid, and thus hinders the subsequent formation of imogolite, leading to the formation of disordered aluminosilicates and pseudoboehmite in the precipitation products. The soluble products formed in the presence of citric acid were characterized by the predominance of proto-imogolite sol and Al-citrate chelates.

In the absence of citric acid, heating the solution containing hydroxy-Al ions and orthosilicic acid to 96-100 °C for 110 h generated a colloidal precipitate (Table 1). The interaction of hydroxy-Al ions with orthosilicic acid caused a marked fall in

pH. In the presence of citric acid, all the heated solutions similarly decreased in pH. However, the pH shift decreased as the citric acid/Al molar ratio of the parent solution was increased (Table 1).

The precipitation of the solid phase products collected by a cellulose nitrate membrane filter (0.01 µm pore size) was invariably hindered by the presence of citric acid. The per cent Si and Al removed from the solution decreased markedly with increase of the citric acid/Al molar ratio (Table 1). In the absence of citric acid, the precipitation product has a $\text{SiO}_2/\text{Al}_2\text{O}_3$ molar ratio of 0.81. However, this ratio was slightly lower than that of the natural imogolite, because of the coexistence of boehmite and bayerite with imogolite in the precipitation product, as discussed below. The $\text{SiO}_2/\text{Al}_2\text{O}_3$ molar ratio of the precipitation products was greatly influenced by the presence of citric acid and also decreased markedly when the citric acid/Al molar ratios of the parent solution was increased from 0.01 to 0.1 (Table 1). The organic C content of the precipitation products increased with increasing the citric acid/Al molar ratio of the parent solution (Table 1). These results show that a part of the citric acid added in the reaction system is complexed with aluminosilicates during their formation at citric acid/Al molar ratios of 0.01-0.05.

At a citric acid/Al molar ratio of 0.1, no precipitates were collected by the ultrafiltration. The IR spectrum of the soluble products obtained by freeze-drying the filtrate at 0.1 citric acid/Al ratio showed absorption maxima at 1,640, 1,445, 1,400, 973, 565 and 343 cm^{-1} and a shoulder or weak absorption at 1,200, 1,080, 860, 800, 490 and 430 cm^{-1} (Fig. 1). The IR data indicate that proto-imogolite (973, 565, 430 and 343 cm^{-1}), silica gel (1,200, 1,080, 860, 800 and 490 cm^{-1}) and citrate (1,640, 1,445 and 1,400 cm^{-1}) are dominant in the soluble products. Electron micrographs of the products after freeze drying showed the presence of irregular, circular and elliptical plates with diameters of 1-20 µm. However, imogolite threads were not identified in the freeze-dried products from the filtrates. The irregular, circular and elliptical silica microplates¹⁵ may be formed from the polymerization process of silicic acid by freeze drying. Therefore, it is obvious that the soluble

* Permanent address: Faculty of Agriculture, Iwate University, Morioka, Japan 020.

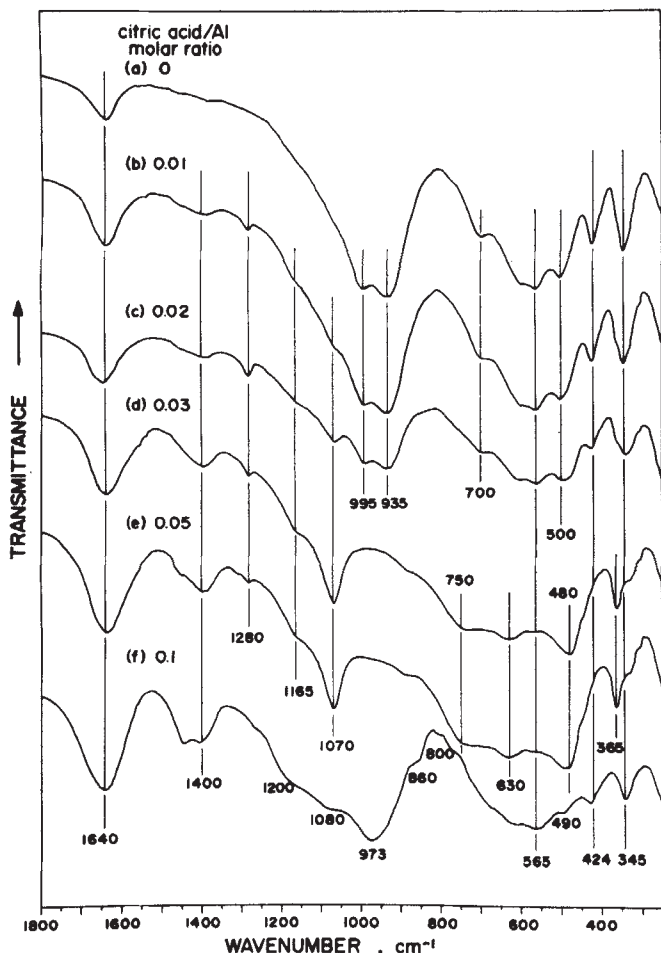


Fig. 1 IR spectra of the colloidal precipitates and the soluble products formed at an initial Si concentration of 1.6×10^{-3} M, Si/Al molar ratio of 0.5, OH/Al molar ratio of 2.0, and citric acid/Al molar ratios of 0–0.1. *a–e*, The precipitates, (*f*), the soluble product. KBr (170 mg) which was dehydrated by heating at 125 °C was ground with all freeze-dried samples (1.0 mg) and the mixtures were pressed to give a 13 mm disk. Spectra were recorded on a Perkin-Elmer 983 spectrometer.

products formed in the presence of citric acid are characterized by the predominance of proto-imogolite^{11,12,14} and soluble Al-citrate chelates. Proto-imogolite is considered to consist of small fragments of imogolite structure¹⁴ with dimensions <100 Å.

The X-ray diffraction pattern of the precipitation products formed at citric acid/Al molar ratios of 0 and 0.01 exhibited a sharp peak at 21.6 Å and a series of broad peaks at 15.1, 8.7 and 6.4 Å. These d spacings are attributed to the presence of imogolite^{11,13}. Small amounts of boehmite and bayerite, which gave X-ray diffraction peaks at 6.1 and 4.7 Å, were found in the precipitation product formed from the solution in the absence of citric acid. At citric acid/Al molar ratios of 0.02–0.05, however, the precipitation products were poorly crystalline to noncrystalline to X ray.

Figures 1 and 2 show respectively the IR spectra and electron micrographs of the synthetic products. The IR spectrum of the precipitation product formed in the absence of citric acid (Fig. 1) resembles that obtained from naturally-occurring imogolite^{2,4,8} and synthetic imogolite^{11–13}. The absorption maxima of the precipitation product resolves at 995 and 935 cm^{-1} , which are mainly due to the Si–O and/or Si(Al)–O stretching vibrations of synthesized aluminosilicates. The IR spectrum of the precipitation product also indicates a distinct absorption band at 345 cm^{-1} and its intensity is correlated with

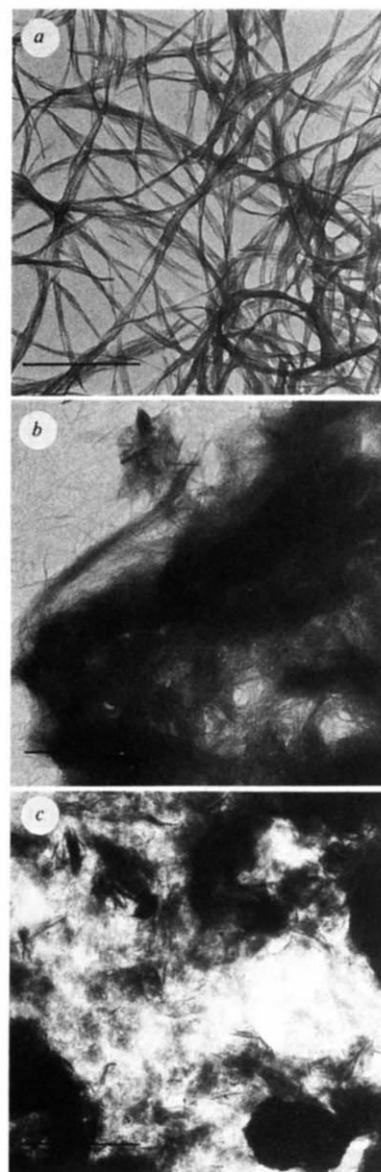


Fig. 2 Transmission electron micrographs of the colloidal products formed at an initial Si concentration of 1.6×10^{-3} M, Si/Al molar ratio of 0.5 and OH/Al molar ratio of 2.0. *a*, Citric acid/Al molar ratio = 0; *b*, citric acid/Al molar ratio = 0.02; *c*, citric acid/Al molar ratio = 0.03. Scale bar, 0.5 μm .

that of an absorption band at 424 cm^{-1} . These features are characteristic of natural imogolite^{2,3,16}. The electron micrograph of the precipitation product formed from the solution in the absence of citric acid shows the dominance of imogolite threads, which appear to have diameters of 20–300 Å (Fig. 2*a*). The high magnification electron micrograph of the precipitation product showed that the imogolite threads consist of finer filiform units, which run parallel with a diameter of about 20 Å, and that hollow allophane spherules were not identified in the precipitate. In the presence of citric acid, the characteristic absorption bands of imogolite at 1,000–900 and 345 cm^{-1} were distorted on increasing the citric acid/Al molar ratio of the parent solution from 0.01 to 0.05 (Fig. 1).

At a citric acid/Al molar ratio of 0.01, imogolite predominated in the precipitation products (Table 1). However, the precipitation products formed from the solution at 0.02 citric acid/Al molar ratio were poorly crystalline to X ray (Table 1) and their absorption bands at 995, 935 and 345 cm^{-1} were significantly distorted (Fig. 1). Furthermore, the tubular morphology of imogolite was considerably distorted (Fig. 2*b*).

Table 1 Influence of citric acid on the interaction of hydroxyl-Al ions with orthosilicic acid

Citric acid/Al molar ratio (parent solution)	$pH_{(0)}$	$pH_{(1)}$	% Si and Al removed from solution		SiO_2/Al_2O_3 molar ratio of the precipitation products	% Organic carbon content in the precipitation products	mol of citrate kg^{-1} of the precipitation products	% Imogolite content in the precipitation products	Mineral composition of the precipitation products
			Si	Al					
0	4.36	2.96	52.2	64.3	0.81	0	0	96	Im, Bm, By,
0.01	4.36	3.06	35.4	52.0	0.67	2.49	0.35	73	Im, X, Ps
0.02	4.35	3.07	12.5	31.9	0.43	3.78	0.52	32	(Im), X, Ps
0.03	4.35	3.11	4.4	17.5	0.27	4.62	0.64	0	Ps, X
0.05	4.36	3.17	0.6	4.4	0.14	8.80	1.22	0	Ps, X
0.1	4.36	3.24	0	0	0	—	—	—	NP

Orthosilicic acid, which was prepared by passing Na-metasilicate through a H^+ -saturated Dowex 50W-X8 cation exchange resin column^{24,25}, and citric acid were simultaneously mixed with $AlCl_3 \cdot 6H_2O$ solution to give the final Si concentration of 1.6×10^{-3} M, the Si/Al ratio of 0.5 and the citric acid/Al ratios of 0–0.1. About 900 ml of the solutions were titrated with 0.1 M NaOH with continuous stirring at the rate of ~0.5 ml NaOH per min to an OH/Al ratio of 2.0. The solutions containing citric acid were adjusted with 0.1 M NaOH to the same pH conditions as the solution without citric acid to avoid acidity effects. After dilution to 1,000 ml, about 900 ml of the solution was heated at 96–100 °C for 110 h. The precipitation products were collected through ultrafiltration using a cellulose nitrate membrane filter of 0.01 μm pore size. The total Al and Si in the parent solution and the filtrate were determined^{26,27}. The per cent Si and Al removed from solution were respectively determined by taking the difference between the Si and Al concentrations in the parent solution and those in the filtrate. $pH_{(0)}$, pH of parent solution after addition of NaOH; $pH_{(1)}$, pH of suspension after 110 h heating at 96–100 °C. Carbon content in the precipitation products was determined by using the CHN analyser. Imogolite content in the precipitation products was determined on the basis of the absorbance at near 345 cm^{-1} in the IR spectra¹⁶. Imogolite separated from the Kitakami pumice bed was used as standard^{28,29}. Bm, boehmite; By, bayerite; Im, imogolite; (Im), poorly ordered imogolite; Ps, pseudoboehmite; X, unknown disordered aluminosilicates; NP, precipitate ($>0.01 \mu m$) not detectable.

At citric acid/Al molar ratios of 0.03 and 0.05, the precipitation products ranged from poorly crystalline to noncrystalline to X ray (Table 1) and their IR spectra showed no characteristic absorption bands of imogolite (Fig. 1). Instead, the IR spectra of the precipitation products showed the strong absorption maxima at 1,070, 750–480 and 365 cm^{-1} and a shoulder at 1,165 cm^{-1} (Fig. 1) which are characteristic of boehmite^{16,17}. Opaline silica also has a strong absorption at 1,070–1,090 cm^{-1} , with a shoulder at about 1,200 cm^{-1} and a weak absorption at about 800 cm^{-1} (refs 18, 19). However, laminar particles with unique shapes such as circular, elliptical and rhombic forms of opaline silica which are often separated from the A_1 horizon of young Ando soils^{18,19} were not identified in the precipitation products. The morphology of the precipitation products was characterized by irregularly shaped particles, their high electron density aggregates, and cloud-like amorphous materials (Fig. 2c), and was distinctly different from that of boehmite^{13,20}. The SiO_2/Al_2O_3 molar ratio of the precipitation products was notably lower than those of imogolite and opaline silica (Table 1). Therefore, the absorption bands at 1,070 and 1,165 cm^{-1} (Fig. 1) are attributed to pseudoboehmite^{21,22}.

The absorption bands at 1,640 and 1,400 cm^{-1} (Fig. 1) are due to the COO^- bending vibrations of citrate²³. The former band is also due to the HOH bending vibration of adsorbed water. The intensity of the absorption bands at 1,640 and 1,400 cm^{-1} increased with increase in the citric acid/Al molar ratio of the parent solution. Citric acid added to the synthetic

system was largely converted from un-ionized COOH groups to the ionic carboxylate form, as indicated by the appearance of strong absorption bands at 1,640 and 1,400 cm^{-1} (Fig. 1) but not at 1,725 cm^{-1} ($C=O$ of COOH) and 1,200 cm^{-1} ($C-O$ stretch for OH-deformation of COOH) (ref. 23). The weak absorption band at 1,280 cm^{-1} is probably due to the un-ionized COOH groups of citrate complexed with aluminosilicates during precipitation. The IR spectra indicate that citric acid reacts with hydroxy-aluminosilicates and/or hydroxy-Al ions to form COO^- ions. Table 1 shows the imogolite content of the precipitation products. The imogolite content in the products decreased markedly with increase in the citric acid/Al molar ratio and imogolite formation was completely disrupted at citric acid/Al molar ratios of 0.03–0.1.

Our data reveal that citric acid, which is present in terrestrial and aquatic environments, greatly perturbs the interaction of hydroxy-Al ions with orthosilicic acid and thus impedes the subsequent formation of imogolite. The present results thus indicate that low molecular weight organic ligands merit close attention as a means of understanding the formation of imogolite in soils and sediments.

This study was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. We thank the Ministry of Education of Japan for providing a travelling fellowship to K.I., and Professor M. Yoshida, Iwate University, for help in the collection of the pyroclastic sample from the Kitakami pumice bed.

Received 13 July; accepted 25 November 1983.

1. Cradwick, P. D. G. *et al. Nature phys. Sci.* **240**, 187–189 (1972).
2. Wada, K. & Harward, M. E. *Adv. Agron.* **26**, 211–260 (1974).
3. Wada, K. in *Minerals in Soil Environments* (eds Dixon, J. B. & Weed, S. B.) 603–638 (Soil Sci. Soc. Am., Madison, 1977).
4. Wada, K. in *Soils with Variable Charge* (ed. Theng, B. K. G.) 87–107 (New Zealand Soc. Soil Sci., Palmerston North, 1980).
5. Tait, J. M., Yoshinaga, N. & Mitchell, B. D. *Soil Sci. Pl. Nutr., Tokyo* **24**, 145–151 (1978).
6. Ross, G. J. & Kodama, H. *Clays Clay Miner.* **27**, 297–300 (1979).
7. McKeague, J. A. & Kodama, H. *Geoderma* **25**, 189–197 (1981).
8. Farmer, V. C., Russell, J. D. & Berrrow, M. L. *J. Soil Sci.* **31**, 673–684 (1980).
9. Anderson, H. A. *et al. J. Soil Sci.* **33**, 125–136 (1982).
10. Inoue, T. & Wada, K. *Clay Sci., Tokyo* **4**, 61–70 (1971).
11. Farmer, V. C. & Fraser, A. R. *Int. Clay Conf. 1978* (eds Mortland, M. M. & Farmer, V. C.) 547–553 (Elsevier, Amsterdam, 1979).
12. Farmer, V. C., Fraser, A. R. & Tait, J. M. *Geochim. cosmochim. Acta* **43**, 1417–1420 (1979).
13. Wada, S.-I., Eto, A. & Wada, K. *J. Soil Sci.* **30**, 347–355 (1979).

14. Farmer, V. C. in *Colloques int. Cent. natn. Rech. Scient.* **303**, 275–279 (1981).
15. Wada, S.-I. & Nagasato, A. *Soil Sci. Pl. Nutr. Tokyo* **29**, 93–95 (1983).
16. Farmer, V. C., Fraser, A. R., Russell, J. D. & Yoshinaga, N. *Clay Miner.* **12**, 55–57 (1977).
17. Tettenhorst, R. & Hofmann, A. *Clays Clay Miner.* **28**, 373–380 (1980).
18. Shoji, S. & Masui, J. *J. Soil Sci.* **22**, 101–108 (1971).
19. Tokashiki, Y. & Wada, K. *Geoderma* **14**, 47–62 (1975).
20. Beutelspacher, H. & van der Marel, H. W. *Atlas of Electron Microscopy of Clay Minerals Hand their Admixtures*, 201–211 (Elsevier, Amsterdam, 1968).
21. Kodama, H. & Schnitzer, M. *Geoderma* **24**, 195–205 (1980).
22. Violante, A. & Huang, P. M. A. *Meet. Soil Sci. Soc. Am., Agron. Abstr.* p. 284 (Anaheim, California, 1982).
23. Schnitzer, M. in *Soil Organic Matter* (eds Schnitzer, M. & Khan, S. U.) 1–64 (Elsevier, Amsterdam, 1978).
24. Alexander, G. B. *J. Am. Chem. Soc.* **75**, 2887–2888 (1953).
25. Luciu, G. M. & Huang, P. M. *Proc. Soil Sci. Soc. Am.* **38**, 235–244 (1974).
26. Hsu, P. H. *Soil Sci.* **96**, 230–238 (1963).
27. Weaver, R. M., Syers, J. K. & Jackson, M. L. *Proc. Soil Sci. Soc. Am.* **32**, 497–501 (1968).
28. Yoshinaga, N. & Aomine, S. *Soil Sci. Pl. Nutr. Tokyo* **8**(3), 22–29 (1962).
29. Henmi, T. & Wada, K. *Am. Mineral.* **61**, 379–390 (1976).

'Insight' in the pigeon: antecedents and determinants of an intelligent performance

R. Epstein*, C. E. Kirshnit, R. P. Lanza & L. C. Rubin

Department of Psychology and Social Relations, Harvard University, Cambridge, Massachusetts 02138, USA

In 1917 Wolfgang Köhler reported some rather extraordinary instances of problem solving by a number of chimpanzees¹, and his observations have been the subject of controversy ever since^{2,3}. The period of quiescence that sometimes preceded the solution, its sudden onset, and its smooth, continuous emergence were proffered as evidence that (1) contrary to suggestions of learning theorists of the day, problem solving was not necessarily a trial-and-error process, and (2) constructs such as 'insight' were necessary for an adequate account^{1,4-6}. Here, in an attempt to shed further light on these issues, we have replicated with pigeons a classic problem with that Köhler confronted his chimpanzees. Pigeons that had acquired relevant skills solved the problem in a remarkably chimpanzee-like (and, perforce, human-like) fashion. The possible contributions of different experiences were determined by varying the training histories of different birds. We offer a tentative moment-to-moment account of a successful performance.

Köhler placed a banana out of reach in one corner of a room and a small wooden crate about 2.5 m from the position on the floor beneath it. After a number of fruitless attempts by all six chimpanzees in the room to jump for the banana, one of them paced for several minutes, then suddenly moved the box half a metre from the position of the banana "and springing upwards with all his force, tore down the banana"¹. Both research⁷ and theory⁸ suggest that chimpanzees will not solve problems of this sort if they have not first had certain experiences. We speculated that two behaviours had to have been acquired: pushing objects towards targets and climbing on objects to reach other objects. Since a pigeon normally does neither, it seemed an ideal candidate to test the contribution that previous learning might make to success in this problem.

Eleven adult male pigeons served as subjects. Each was maintained at about 80% of the weight it would achieve given free access to food. Most had had a variety of laboratory experience, but none had ever been used in a problem solving experiment. Birds 269WP and 270WP were racing Homers; the others were white Carneau. All sessions were conducted in a cylindrical wire-mesh chamber 69 cm in diameter, except those of birds 110YP, 233WP and 274WP, which were conducted in smaller rectangular chambers. A cardboard box, 8 cm high and with a base 10 cm², was used in some conditions, as was a small facsimile of a banana, 7 cm in length. A standard grain dispenser was attached to the base of each chamber as shown in Fig. 1.

The following history yielded successful performances with all of the birds we tested: (1) A repertoire of 'directional pushing' was established. Each bird was trained to push the box towards a green spot, 4 cm in diameter, which was placed at random positions along the base of the chamber wall(s). Pushing was extinguished in the absence of the green spot. Major training steps included reinforcing aimless pushes; reinforcing pecks to the spot; reinforcing sighting the spot and pushing the box towards it, with the movement of the box constrained by a thin wire; reinforcing sight-and-push behaviour with the wire removed and the box close to the spot; and gradually increasing the distance between the box and the spot⁹. Proficient performances were established in 8, 1 and 4 weeks, respectively, for the

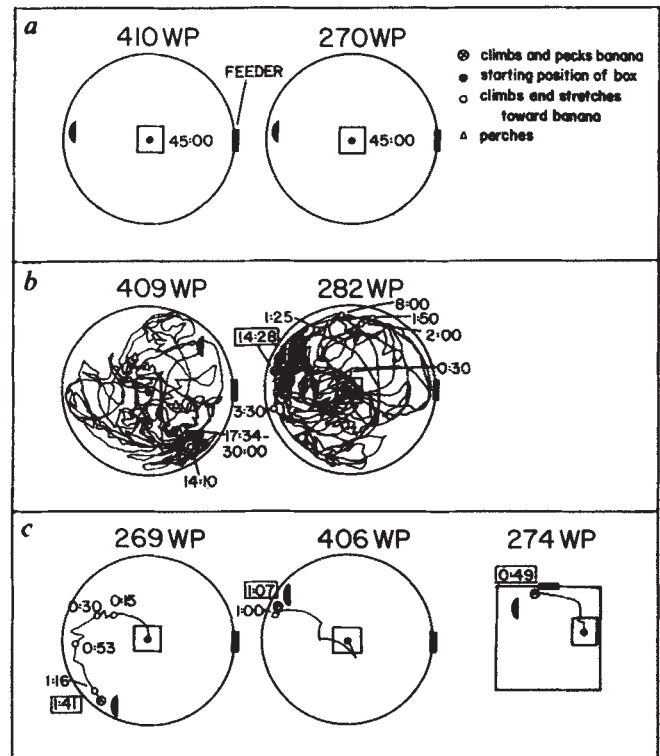


Fig. 1 Birds that had been trained to climb and peck but never to push did not push the box in the test situation (a). Birds that had been trained (i) to climb and peck and (ii) to push the box aimlessly for long periods of time pushed the box over much of the floor space of the chamber. The birds rarely looked up while pushing. One of the birds stopped pushing in the appropriate place and climbed and pecked the banana after having pushed for more than 14 min (b). Birds that had been trained (i) to climb and peck and (ii) to push the box towards a green spot placed at random positions along the base of the chamber solved the problem efficiently and in a manner suggestive of human problem-solving behaviour (c). Other controls are described in the text. The times given are in minutes and seconds. A boxed time is the time to solution.

subjects whose performances are shown in Fig. 1c. The banana was never present during this training. (2) Concurrently, each bird was trained to climb onto the box and peck the banana, which was suspended overhead. The box was fixed in place during this condition, and pecking it was never reinforced. The position of box and banana was changed repeatedly. In the presence of box and banana, the bird would reliably climb onto the box and peck the banana. In the absence of the banana and in the presence of the spot, the bird would push the box towards the spot. (3) Each bird was occasionally placed alone with the banana until the bird neither flew nor jumped towards it.

The following test situation was arranged. The banana was suspended out of reach (41 cm from the floor) at a point (determined by a random number) near an edge of the chamber, and the box was placed elsewhere in the chamber. All test sessions for these and all other subjects were filmed or videotaped.

The performances for the first three subjects were remarkably similar. At first each pigeon appeared to be 'confused'; it stretched and turned beneath the banana, looked back and forth from banana to box, and so on. Then each subject began rather suddenly to push the box in what was clearly the direction of the banana (Fig. 1c). Each subject sighted the banana as it pushed and readjusted the box as necessary to move it towards the banana. Each subject stopped pushing in the appropriate place, climbed and pecked the banana.

A fourth bird (233WP) solved the problem after 24 min. The performance was disrupted by 1,000 W of lighting which had been added to facilitate filming. When, after 20 min, the lighting was reduced, the bird solved the problem in just under 4 min.

* Present address: Cambridge Center for Behavioral Studies, 11 Ware Street, Cambridge, Massachusetts 02138, USA.

We conducted four variations on this training with other pigeons. Two birds (294WP and 273WP) were trained to peck the banana but not to climb. Jumping and flying were extinguished, and the birds were placed alone with the box until they showed no signs of discomfort in its presence. Then the banana was suspended out of reach above it. Each bird stretched repeatedly towards the banana at first. Subject 273WP stumbled onto the box and then fell off. After the first few minutes of each session, attempts to reach the banana ceased. We terminated each session after 10 min. We concluded that the establishment of climbing was probably critical to the solution.

Two birds (270WP and 410WP) were trained to climb and peck but not to push. Jumping and flying were extinguished. Neither bird pushed the box when given the test (Fig. 1a). Two birds (409WP and 282WP) were trained to climb and peck and to push the box around the chamber for long periods of time. They were never trained to push towards a target, nor to push in straight lines. Jumping and flying were extinguished. The birds pushed apparently aimlessly when given the test (Fig. 1b). We concluded that a repertoire of directional pushing was probably critical to an efficient solution.

With one bird (110YP) we established directional pushing and climbing-and-pecking but did not extinguish brute force attempts to reach the banana. Like Köhler's chimpanzee, the bird jumped and flew repeatedly towards the banana for several minutes, then pushed the box towards the banana, climbed and pecked. The solution appeared after about 7 min.

Based on these and other experiments, a tentative, moment-to-moment account of a successful performance can be given. At first stimuli were present which controlled both behaviour with respect to the banana and behaviour with respect to the box. The behaviour we interpreted as a sign of perplexity was probably the result of competition between these behaviours. Behaviour with respect to the banana quickly disappeared, probably because of the recent history of extinction of jumping and flying when the banana was out of reach (compare the performance of bird 110YP). The birds may have begun to push because, as behaviour with respect to the box increased in relative frequency, the birds faced the box more directly, which was very nearly the stimulus in the presence of which pushing had been reinforced (the green spot was absent). Why the animals pushed towards the banana is unclear and still under investigation. A process similar to what some call 'functional generalization'¹⁰ (as opposed to generalization based solely on common physical characteristics) seems to be involved. Birds that were trained to push towards the spot but not to peck the banana did not push towards the banana in the test situation but did push towards the banana when subsequently trained to peck it. In other words, the birds pushed towards the banana apparently for the 'right reasons'—because they had learned directional pushing and because some history of reinforcement had made the banana 'important'. Directional performances may also have been produced by a summation of prevailing responses: banana-directed pecks may have strengthened banana-directed pushes (N. E. Miller, personal communication). The birds stopped pushing in the right place because of a phenomenon called 'automatic chaining': in the course of pushing towards the banana, they set up for themselves a stimulus (box-under-banana) that controlled other behaviour (climbing and pecking).

We appear to have in hand an instance of 'insightful' problem solving. The suddenness, directness and continuousness of the performances satisfy Köhler's criteria for 'genuine' or 'insightful' solutions^{1,11}, and people viewing the tapes have liberally attributed a wide range of human emotions and thoughts to the pigeons. A surprisingly common comment was, "Did the pigeon really do that?" We may also have in hand an account of similar performances in chimpanzees and children, for the experiences we provided are ones that they have probably had before they are successful in similar situations, and the behavioural processes we have invoked are fairly general in the animal kingdom¹².

We emphasize that we did not train the birds to push the box towards the banana; that, except during very early stages of training, behaviour with respect to the box was never reinforced in the absence of the green spot and that such behaviour was deliberately extinguished; that pushing the box was never reinforced in the presence of the banana and that such behaviour was deliberately extinguished; and that the spot was absent during the test. The successful performances must consequently be regarded as genuinely novel.

The work was supported in part by a NSF grant to Harvard University and a NIH grant to the Foundation for Research on the Nervous System. We thank J. Cerella, J. E. Mazur and B. F. Skinner for suggestions and E. Stoddard for assistance. The data reported here are based on an analysis of videotapes and films by an independent observer. Some preliminary results were reported elsewhere¹² before the tapes and films had been examined.

Received 23 August; accepted 8 December 1983.

1. Köhler, W. *The Mentality of Apes* (Routledge & Kegan Paul, London, 1925).
2. Chance, M. R. A. *Man* **60**, 130–135 (1960).
3. Weisberg, R. & Alba, J. J. *exp. Psychol.: Gen.* **110**, 169–192 (1981).
4. Maier, N. R. F. *J. comp. Psychol.* **12**, 181–194 (1931).
5. Duncker, K. *Psychol. Monogr.* No. 270 (1945).
6. Ellen, P. J. *exp. Psychol.: Gen.* **111**, 316–325 (1982).
7. Birch, H. G. J. *comp. Psychol.* **38**, 367–383 (1945).
8. Hull, C. L. *Psychol. Rev.* **42**, 219–245 (1935).
9. Epstein, R. & Medalie, S. D. *Behav. Analys. Lett.* **3**, 241–247 (1983).
10. Bruner, J. S., Goodnow, J. J. & Austin, G. A. *A Study of Thinking* (Wiley, New York, 1956).
11. Koffka, K. *The Growth of the Mind* (Kegan Paul, London, 1924).
12. Epstein, R. *Behav. Analyst* **4**, 43–55 (1981).

Learning and memory mutations impair acoustic priming of mating behaviour in *Drosophila*

Charalambos P. Kyriacou

Department of Psychology, University of Edinburgh,
Edinburgh EH8 9TJ, UK

Jeffrey C. Hall

Department of Biology, Brandeis University, Waltham,
Massachusetts 02254, USA

The courtship song in *Drosophila melanogaster* has two components, a low-frequency hum and a train of pulses with a species-specific interpulse interval (IPI) of 30–40 ms^{1,2}. The IPIs oscillate rhythmically, with periods between 50 and 60 s in wild-type males³. When females are stimulated with artificial songs in the presence of courting but silent (wingless) males, the 'pulse song' and its oscillation can enhance mating success^{4–6}. If separated males and females are first simultaneously primed with invariant 34-ms IPIs, their subsequent mating success is improved⁷. However, exclusive prestimulation of females leads to faster mating only when the hum component of the song is applied, not constant 34-ms IPIs⁵. We have re-examined these findings by testing whether prior exposure of females to a rhythmic pulse song speeds up subsequent mating performance. We report here that it does. Furthermore, learning and memory mutations⁸, expressed in the females to whom songs are being played, either 'block' or attenuate the effectiveness of acoustical priming.

We used an electronic song simulator to prime virgin 'Canton-S' females with 2 min of artificial song⁶. Three types of pulse song were presented, one which consisted of an invariant 35-ms IPI, and two rhythmic songs which oscillated about this 35-ms value but were programmed with either a 35- or 55-s period. These two rhythms are characteristic of *Drosophila simulans* and *Drosophila melanogaster* songs respectively³. In other tests

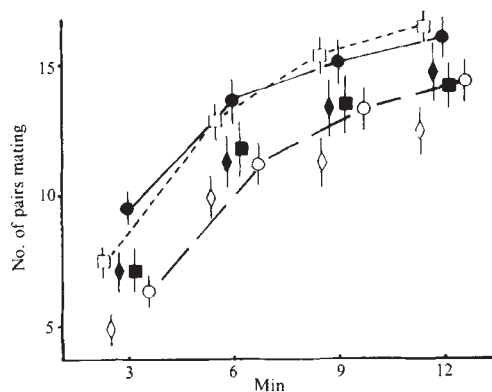


Fig. 1 Mating speed of wild-type (Canton-S) females prestimulated with artificial courtship songs. Twenty-five virgin 4–5-day-old females were placed in a $5 \times 5 \times 1$ cm chamber 7 cm above a loudspeaker and given 3 min to accustom themselves to their new environment. Two minutes of song were played to the females at an intensity level of 97 dB, in an intermittent pattern of 2 s song/3 s silence (see ref. 5). At the end of the 2-min prestimulation, an identical container containing 25 similarly aged Canton-S males was placed above the females' chamber, and the partitions between the two removed, allowing the males and females to mix. The cumulative number of flies mating at 1 min intervals was recorded; 15 replications of each type of prestimulation were performed; experiments were carried out at 25 °C. ●, 55-s rhythmic pulse song (mean IPI = 35 ms, amplitude = + and -10%); ◆, 35-s rhythmic pulse song (mean IPI = 35 ms, amplitude = + and -10%); ■, invariant 35-ms IPI; □, hum song (160 Hz); ○, silence; ◇, white noise (90 dB).

the females were prestimulated with hum song, white noise or were left in silence. After 2 min of prestimulation, 25 Canton-S males were added to the females, and the number of flies mating was recorded at 1-min intervals. Analysis of variance with respect to the cumulative number of pairs mated after 3 min revealed a highly significant difference between the effects of the various preconditions ($F = 3.94$, $P < 0.01$). Figure 1 illustrates the results. All treatments, including silence, led to significantly more matings than white noise ($P < 0.05$, Newman-Keuls procedure⁹); we therefore used the response of females previously left in silence as our control. We found that prior exposure to the 55-s rhythmic oscillation was the only pretreatment which led to significantly faster mating than this control ($P < 0.05$). Later on, at 6, 9 and 12 min after introduction of the males, the effectiveness of the hum song equalled that of the rhythmic pulse song.

Females therefore appear to 'store' rhythmic 55-s pulse songs, in that their receptivity to subsequent mating attempts is enhanced. The relative ineffectiveness of a 35-s 'simulans-type' rhythm may indicate a species specificity in this priming effect (see ref. 6). Hum song can also effectively prestimulate females, confirming Schilcher's results⁵.

Several X-linked mutations in *D. melanogaster* disrupt the fly's ability to associate a chemical odorant with electric shock⁸. Two mutants, *dunce* (*dnc*) and *rutabaga* (*rut*), forget so rapidly in most tests of associative conditioning that no training effects can be revealed^{10–13}. Another mutant, *amnesiac* (*amn*), associates specific odours with the aversive stimulus, but retains the effects of training for a much shorter time than the wild type^{13,14}. We tested mutant and normal females for acoustic priming using three pretreatments—55-s rhythmic pulse song, hum song and silence. We also extended the experiment to obtain decay curves for the after-effects by introducing the males to the females either immediately (time 0) or with delays of 1, 3 or 5 min after the end of the song stimulus.

Figure 2 reveals that females from a Princeton subline of Canton-S have the usual enhanced receptivity after prestimulation by rhythmic pulse song or by the hums, when males are

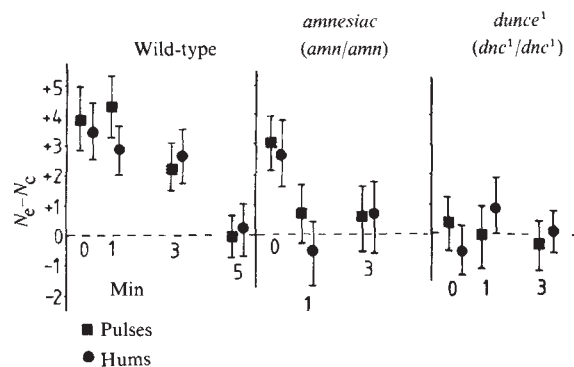


Fig. 2 Mating speed decay curves for a wild-type Canton-S subline and mutant *dunce*¹ or *amnesiac* females, prestimulated with 55-s rhythmic pulse song (squares) or hum song (circles), compared with silence control. The mutant stocks had been made nearly isogenic with this Canton-S subline and all strains were obtained from Dr W. Quinn, Princeton University. In these and all subsequent experiments, songs were played to the females at a reduced level of 90 dB. Males were presented to the females immediately after the 2-min prestimulation (time 0) or after delays of 1, 3 or 5 min. The cumulative number of matings was recorded every minute until the 5th minute (no matings can be missed as the duration of copulation is ~20 min in this species). Analysis of variance was performed on the numbers mating after 5 min. Several sets of experiments with three matched trials per set (silence, pulses, hums) were run for each female genotype. The number mating at the 5th minute of the appropriate silence control condition (N_c) was then subtracted from the number mating in each of the two corresponding experimental trials (N_e). These values ($N_e - N_c$) were then plotted as means $\pm$ s.e.m. for the 15–20 sets of trials that were run for each delay condition (that is, at times 0, 1, 3 and 5 min).

added immediately (time 0) or after a 1-min delay ($P < 0.05$, except for hum song at 1 min which just fails to reach significance). These enhancing after-effects decay when a 3-min delay is imposed, but are absent after a 5-min lag. In contrast, females homozygous for *dnc*¹ and *dnc*² do not mate faster after prestimulation (Figs 2, 3). Thus, two independently isolated *dnc* mutations give the same behavioural defect, suggesting that these results are due to genetic changes at this locus, not to differences in genetic background. Figure 3 also reveals that a further group of females (*FM7(dnc⁺)/+*) were strongly primed when males were added immediately after the song stimulus ($P < 0.01$) or after a delay of 1 min ($P < 0.05$). However, coisogenic heterozygous *dnc*²/+ females give a significant prestimulation effect for both songs only at time 0 ($P < 0.01$). The effects with *dnc*²/+ females are similar to those induced with the *amn* memory mutation. *Amn/amn* females also reveal a priming effect ($P > 0.05$) but only when males were presented immediately (time 0) after the song stimuli (Fig. 2).

In tests involving the *rutabaga* gene, homozygous (*rut/rut*) and hemizygous (*Df/rut*) females exhibit the same behavioural phenotype as *dnc* females, with no after-effects observed (Fig. 4). The results from the tests of hemizygous *Df/rut* females strongly suggest that this impairment maps to the *rut* locus. Heterozygosity for this mutation (*FM7/rut*) or for a deletion of the gene (*FM7/Df*) gives significantly enhanced mating speeds ($P < 0.05$ in both cases), but only when males were mixed with these females immediately after prestimulation (Fig. 4). Therefore, like the *dnc*²/+ or *amn/amn* genotypes (Figs 2, 3), the *rut* mutation or deletion behaves as a 'memory' variant when either is heterozygous with the normal allele. These results with *rut* and *dnc*² heterozygotes are strikingly similar to the effects observed with these flies in associative learning¹².

Our results (Figs 2–4) imply that the original stimulus needs to be reinforced by the tester males' courtship fairly soon after the female first hears the artificial courtship sounds. That is, even though enhanced mating speeds can be detected up to at least 12 min post-treatment (Fig. 1), the after-effects will be observed only if the prestimulated female is courted within

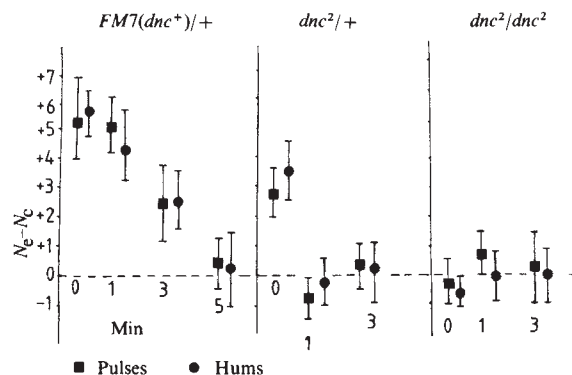


Fig. 3 Mating speed decay curves for *FM7/+* (normal), *dunce*²/*+* or homozygous *dunce*² females, after prestimulation with pulses and hums. We partly isogenized the genetic background of *dunce*² (refs 11, 21) with the Princeton subline of Canton-S, by crossing both normal and mutant strains to flies carrying the multiple inverted X-chromosomal 'balancer' *FM7*, and performing backcrosses for several generations. By crossing *FM7/dunce*² females to males of the Princeton Canton-S wild-type subline, we obtained both the *FM7/+* and *dunce*²/*+* females. See legends to Figs 1 and 2 for additional explanation of the experiments and data presentation.

1–3 min of the prior song exposure. This is the critical variable, not merely the total amount of time elapsed between the end of the prestimulation and the moment of data collection (compare legends to Figs 2 and 3). This effect appears similar to the psychological phenomenon of 'savings'.

We considered whether artefacts could explain our results. First, we asked if *dnc*, *rut* and *amn* females suffered varying degrees of deafness, thus rendering our attempts to prestimulate them ineffective. We took advantage of Schilcher's observation that on hearing pulse songs, males court each other vigorously². We observed that groups of *dnc*², *rut*, *amn* and wild-type males showed a seven- to nine-fold increase in courtship interactions when stimulated with artificial pulse song, whereas genetically deafened males (simultaneously expressing the *aristales* and *thread* mutations³¹) showed no such increase in 'homosexual' courtship. Consequently, males carrying the learning and memory mutations are not deaf, and it seems unlikely that the corresponding mutant females would be impaired in this manner.

Second, inspection of our data for the number of matings per minute rules out the possibility that prestimulated *dnc*¹, *dnc*² or *rut* females were mating faster during the early stages of the observation period. 'Ceiling effects' were also absent, in that none of the mutant females was so highly receptive that prestimulation was unlikely to lead to further enhancements in mating speed.

Third, we also examined whether these learning and memory mutations affected the male *Drosophila*'s courtship song. If this was the case and genetic coupling was operating between the female reception and the male transmission of the song, as in crickets^{15,16}, the mutant females would not respond to artificial songs bearing the wild-type characteristics. We therefore recorded and analysed the songs of *dnc*, *amn* and *rut* males. All three mutant types produced normal pulses and hums. More specifically, both the overall values of IPI and the IPI rhythm were all well within the normal wild-type range (for example, the periods for *dnc*¹ songs were 54.2 ± 2.8 s, $n = 5$; for *amn*, 55.5 ± 2.5 s, $n = 4$; and for *rut*, 56.9 ± 2.3 s, $n = 4$).

We therefore suggest that females expressing conditioning mutations are defective in their ability to be primed by the courtship song because such priming involves mechanisms of information storage and retrieval. These mutants also perform relatively poorly in other learning tasks, such as associative reward conditioning¹³, visual learning¹⁷, courtship conditioning of male-specific behaviours^{18,19} and a form of sensitization

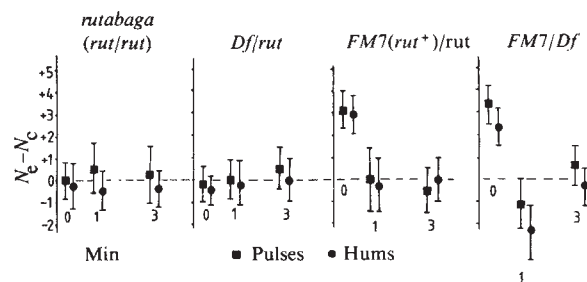


Fig. 4 Mating speed decay curves for prestimulated *rutabaga* females and for similarly treated females carrying various combinations of *rut*, *FM7* (an X-chromosomal balancer), or an X chromosome deleted of the *rut* locus (called *Df(1)KA9*, lacking the X chromosome segment 12E1–13A5 which uncovers *rut*)²². Other pertinent details are given in the legends to Figs 2 and 3.

involving sugar stimuli²⁰. The latter is of particular relevance here, because the acoustic priming of females is similar to a sensitization procedure. The enhanced receptivity of such females can be considered as an increase in responsiveness following the 'strong' stimulation by the artificial courtship songs.

If we have uncovered an example of acoustic sensitization, it is of further interest that both *dnc* and *rut* mutations affect cyclic AMP metabolism^{21–24}. This small molecule is an important 'signalling' component in the control of sensitization of the gill-withdrawal reflex in *Aplysia* (reviewed in ref. 25). Modulation of cyclic AMP levels also appears to be involved in certain types of circadian rhythms^{26–28}. The *dnc* or *rut* mutations in *Drosophila*, however, do not noticeably influence the short-term rhythm in courtship song.

Biometrical analysis of different *Drosophila* strains measured for their learning ability in tests using odours and electric shocks has indicated a strong relationship between this conditioning phenotype and evolutionary fitness²⁹. Variations in sexual behaviour are, of course, also related to fitness. Our demonstration (together with other work^{18,19,30}) that conditioning mutations disrupt the normal storage and retrieval of information during sexual interactions, suggests at least one reason that natural selection might favour behavioural plasticity in *Drosophila* (also see ref. 13). We have also reinforced the view that these mutations affect a central process in the nervous system which regulates associative and non-associative conditioning. An eventual understanding of the cellular defects induced by conditioning mutations, including the role of cyclic AMP and other molecules²⁵, may illuminate the brain mechanisms mediating reproductive behaviour in these insects.

This work was initiated while C.P.K. held a NATO fellowship at Brandeis University, and was completed with the support of a SERC Advanced Fellowship to this investigator. We acknowledge SERC grant GRB 94595 to C.P.K. and USPHS grant GM-21473 to J.C.H. We thank William Quinn and Margaret Livingstone for providing stocks, Angela Reid for technical assistance and Donald Gailey, William Harris, Vincent Henrich and John Ringo for comments on the manuscript.

Received 14 September; accepted 5 December 1983.

- Ewing, A. W. & Bennet-Clark, H. C. *Behaviour* **3**, 287–301 (1968).
- von Schilcher, F. *Anim. Behav.* **24**, 18–26 (1976).
- Kyriacou, C. P. & Hall, J. C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6729–6733 (1980).
- Bennet-Clark, H. C. & Ewing, A. W. *Anim. Behav.* **17**, 755–759 (1969).
- von Schilcher, F. *Anim. Behav.* **24**, 622–625 (1976).
- Kyriacou, C. P. & Hall, J. C. *Anim. Behav.* **30**, 794–801 (1982).
- Bennet-Clark, H. C., Ewing, A. W. & Manning, A. *Behav. Biol.* **8**, 763–769 (1973).
- Quinn, W. G. & Greenspan, R. J. *A. Rev. Neurosci.* **7**, 67–93 (1984).
- Winer, B. J. *Statistical Principles in Experimental Design*, 2nd edn (McGraw-Hill, London, 1971).
- Dudai, Y., Jan, Y.-N., Byers, D., Quinn, W. G. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1684–1688 (1976).
- Dudai, Y. *J. comp. Physiol.* **130**, 271–275 (1979).
- Dudai, Y. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5445–5448 (1983).
- Tempel, B. T., Bonini, N., Dawson, D. R. & Quinn, W. G. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1482–1486 (1983).
- Quinn, W. G., Sziber, P. P. & Booker, R. *Nature* **277**, 212–214 (1979).

15. Alexander, R. D. *Evolution* **16**, 443-467 (1962).
16. Pollack, G. S. & Hoy, R. *Science* **204**, 429-432 (1979).
17. Folkers, E. J. *Insect Physiol.* **28**, 535-539 (1982).
18. Siegel, R. W. & Hall, J. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3430-3434 (1979).
19. Gailey, D. A., Jackson, F. R. & Siegel, R. W. *Genetics* **102**, 771-782 (1982).
20. Duerr, J. S. & Quinn, W. G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3646-3650 (1982).
21. Byers, D., Davis, R. & Kiger, J. A. Jr *Nature* **189**, 79-81 (1981).
22. Livingstone, M. S., Sziber, P. P. & Quinn, W. G. *Cell* (submitted).
23. Davis, R. L. & Kauvar, L. M. *Adv. Cyclic Nucleotide Res.* (in the press).
24. Livingstone, M. S., Sziber, P. P. & Quinn, W. G. *Neurosci. Abstr.* **8**, 384 (1982).
25. Kandel, E. R. & Schwartz, J. H. *Science* **218**, 433-443 (1982).
26. Eskin, A., Corrent, G., Lin, C.-Y. & McAdoo, D. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 660-664 (1982).
27. Eskin, A. & Takahashi, J. S. *Science* **220**, 82-84 (1983).
28. Feldman, J. F. & Dunlap, J. C. *Photochem. Photobiol. Rev.* **7**, 319-368 (1983).
29. Hewitt, J. K., Fulker, D. W. & Hewitt, C. A. *J. comp. Psychol.* **97**, 52-58 (1983).
30. Gailey, D. A., Jackson, F. R. & Siegel, R. W. *Genetics* (in the press).
31. Burnet, B., Eastwood, L. & Connolly, K. *Anim. Behav.* **25**, 460-464 (1977).

Relationships between brain noradrenergic activity and blood glucose

G. A. Smythe, H. S. Grunstein, J. E. Bradshaw, M. V. Nicholson & P. J. Compton

Garvan Institute of Medical Research, St Vincent's Hospital, Darlinghurst, New South Wales 2010, Australia

Glucose is the principal energy substrate for the brain and studies have shown that the brain is able to increase glucose availability in the face of glucose starvation (neuroglycopenia)¹⁻³. The mechanisms, believed to be hypothalamic³, that may be involved in a brain/blood glucose control system have not yet been identified. We have used novel techniques for assessing brain monoamine neuronal activity to investigate its relationship to blood glucose concentrations in the rat. We describe here two important relationships which emerge from these studies. One is that activation of hypothalamic noradrenaline (NA) activity following stress is associated with concurrent increases in plasma glucose concentrations. This relationship is linear and independent of the adrenal or pituitary glands. The second is an inverse relationship between plasma glucose concentration and hypothalamic NA neuronal activity—high blood glucose levels significantly inhibited the hypothalamic NA activity responses to stress, α_2 -adrenergic blockade and adrenalectomy. Thus glucose (or a metabolite of it) seems to provide a negative feedback signal sensed by hypothalamic NA neuronal systems which, in turn, appear to stimulate liver glucose output by a neural mechanism.

Techniques such as computerized gas chromatography/mass spectrometry enable the neuronal activities of the major brain monoamine neurotransmitters dopamine (DA), NA and serotonin (5-hydroxytryptamine, 5-HT) to be assessed from the ratio of the concentration of the major neuronal metabolite of the monoamine to the concentration of the monoamine itself⁴. With concurrent measurement of pituitary hormonal status, this technique has allowed the elucidation of the neuroendocrine control of growth hormone⁵, thyroid stimulating hormone⁶ and adrenocorticotrophic hormone (ACTH) release⁴ in the rat. In investigating the central control of ACTH and corticosterone release following stress in the rat, we demonstrated a highly significant and apparently causal relationship between the hypothalamic ratio of the intraneuronal NA metabolite 3,4-dihydroxyphenylethylene glycol (DHPG) to NA (DHPG/NA ratio) and ACTH and corticosterone levels⁴. Because of the role of glucocorticoids and possibly also β -endorphin⁷ (which is released from the pituitary together with ACTH) in promoting liver glucose output and their possible contribution to the phenomenon of stress-hyperglycaemia, we examined the relationship between hypothalamic NA activity and blood glucose in stressed rats and in rats treated with the α_2 -adrenergic blocker yohimbine, which is a potent stimulus for NA neuronal activity^{4,8,9} and ACTH release⁴.

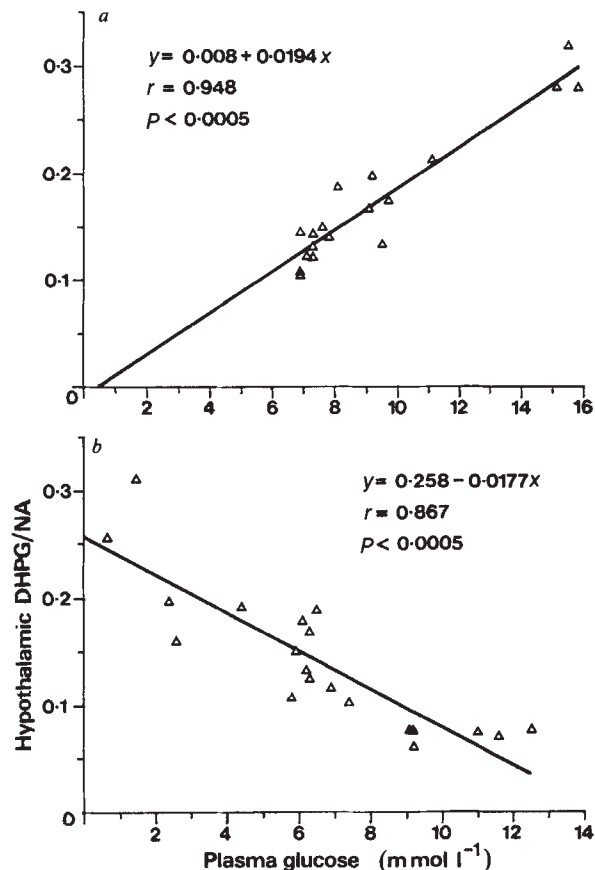


Fig. 1 *a*, The relationship between the medial basal hypothalamic DHPG/NA ratio and plasma glucose in 200 g male Wistar rats 30 min after the induction of neuroglycopenia by 2-DG administration (164-500 mg per kg i.p.). *b*, The relationship between the medial basal hypothalamic DHPG/NA ratio and plasma glucose in 200 g male Wistar rats in different states of glycaemia ranging from hyperglycaemia (plasma glucose >9 mmol l⁻¹, induced by chronic dexamethasone treatment (66 µg per 24 h) in drinking water for 1 day) to hypoglycaemia induced by bilateral adrenalectomy without or with additional treatment with insulin (2 U per kg, 30 min before sampling in animals with plasma glucose <5 mmol l⁻¹).

Plasma glucose concentrations in rats subjected to physiological or chemical stress or yohimbine treatment rose significantly, together with the levels of medial basal hypothalamic NA neuronal activity (assessed from the ratio DHPG/NA), serum ACTH and corticosterone (see Table 1). We also found that within each stress study there was a significant positive correlation between hypothalamic NA neuronal activity and plasma glucose. Stepwise multiple linear regression analysis indicated that the relationship between hypothalamic NA neuronal activity and plasma glucose was independent of the relationship with corticosterone ($F = 59$, $P < 0.0001$). These data thus raised the possibility that the increase in glucose level by hypothalamic NA activity resulted from either direct neural stimulation of the liver or mediation of circulating hormones such as adrenal catecholamines, pancreatic glucagon or pituitary β -endorphin. To test these possibilities, we studied the effects of the administration of 2-deoxy-D-glucose (2-DG) in normal intact rats, and in adrenalectomized and hypophysectomized rats. 2-DG administration is known to cause neuroglycopenia in the central nervous system, by blocking the normal metabolism of glucose, and to result in hepatic glycogenolysis and increased outflow of glucose from the liver².

Thirty minutes after 2-DG administration to intact, adrenalectomized and hypophysectomized rats, hypothalamic NA activity and plasma glucose were markedly stimulated (all $P < 0.0005$) and, similar to the stress experiments, there was a highly

Table 1 The stimulatory effects of various stresses and the α_2 -antagonist, yohimbine, on medial basal hypothalamic NA status and plasma concentrations of ACTH, corticosterone and glucose in 220 g male Wistar rats

Stress	Hypothalamic DHPG/NA	Serum ACTH (pg ml ⁻¹)	Serum corticosterone (nmol l ⁻¹)	Plasma glucose (mmol l ⁻¹)
Controls	0.09 ± 0.004	—	184 ± 39	6.1 ± 0.1
Exercise*	0.22 ± 0.01	—	1,343 ± 100	7.1 ± 0.4
Controls	0.102 ± 0.004	84 ± 13	230 ± 60	7.4 ± 0.1
Cold†	0.21 ± 0.01	230 ± 60	1,270 ± 120	10.0 ± 0.2
Controls	0.09 ± 0.006	96 ± 14	112 ± 32	6.3 ± 0.2
Ether‡	0.14 ± 0.01	382 ± 48	1,014 ± 87	8.9 ± 0.3
Controls	0.10 ± 0.005	155 ± 46	259 ± 79	6.4 ± 0.1
Morphine withdrawal§	0.31 ± 0.003	350 ± 32	1,250 ± 57	12.9 ± 1.2
Controls	0.10 ± 0.005	90 ± 8	490 ± 120	7.3 ± 0.1
Yohimbine	0.20 ± 0.02	380 ± 23	1,000 ± 97	8.0 ± 0.1

Values are mean ± s.e.m. There were at least six rats per group. All differences from control values were highly significant ($P < 0.01$) except for the case of glucose level after exercise which was significantly different from the control level ($P < 0.025$).

* Rats killed 20 min after 10 min of running on a treadmill at 19 m min⁻¹.

† Rats killed 20 min after 2 min in ice-water bath.

‡ Rats killed 15 min after 1 min exposure to ether vapour.

§ Morphine-dependent rats killed 30 min after administration of naloxone (1 mg per kg intraperitoneal, i.p.).

|| Rats killed 30 min after administration of the α_2 -antagonist, yohimbine HCl (8 mg per kg i.p.).

significant positive correlation between hypothalamic DHPG/NA and glucose. The relationship for the intact control rats is illustrated in Fig. 1a and the data for adrenalectomized and hypophysectomized rats, where only one dose of 2-DG was given, are shown in Table 2. These data strongly support the concept of direct neural stimulation of liver glucose output as proposed by Shimazu¹⁰. The glucose response to increased hypothalamic NA activity observed in this investigation cannot be primarily mediated by either adrenal catecholamines or glucocorticoids, nor by pituitary hormones. Pancreatic glucagon release could be a factor contributing to the release of glucose seen in these studies, but we found that administration of glucagon to our animals in pharmacological doses failed to increase plasma glucose to the levels achieved after 2-DG treatment; for example, see the data for adrenalectomized rats given in Table 2.

From the above data we hypothesized that the increased hypothalamic NA neuronal activity associated with stress or neuroglycopenia ultimately makes more glucose available for

brain function, therefore we would expect any change in brain glucose availability or metabolism to affect hypothalamic NA activity. We tested the glucose feedback aspects of this hypothesis in two ways. First, in animals in different states of glycaemia ranging from hyperglycaemia (induced by chronic low-dose dexamethasone administration) to hypoglycaemia (induced by adrenalectomy without or with insulin treatment) we observed a highly significant inverse relationship between hypothalamic NA activity and plasma glucose (see Fig. 1b). Second, we examined whether high glucose levels might attenuate the hypothalamic NA responses to stress, α_2 -adrenergic blockade and adrenalectomy. This indeed was the case and, as illustrated in Fig. 2, dextrose administration significantly reduced the elevated NA activity in adrenalectomized rats and also in intact rats subjected to cold-swim stress. Similarly, high endogenous glucose levels in streptozotocin-induced diabetic rats were associated with significantly lower hypothalamic NA activity responses to yohimbine administration, also shown in Fig. 2.

Table 2 Effects of glucose, 2-DG and glucagon administration to adrenalectomized and of 2-DG administration to hypophysectomized male Wistar rats on the hypothalamic DHPG/NA ratio and plasma glucose concentrations

Group	Hypothalamic DHPG/NA	Plasma glucose (mmol l ⁻¹)
ADRX control	0.177 ± 0.009	7.1 ± 0.2
ADRX + dextrose	0.152 ± 0.008*	7.9 ± 0.2*
ADRX + 2-DG	0.29 ± 0.03	11.1 ± 0.8
ADRX + glucagon	0.157 ± 0.005*	9.8 ± 0.3†
HYPOX control	0.151 ± 0.009	6.4 ± 0.3
HYPOX + 2-DG	0.294 ± 0.02†	17.9 ± 1.0†

The animals were bilaterally adrenalectomized (ADRX), then maintained on 0.9% saline instead of normal drinking water until the experiment 5 days later. Rats were hypophysectomized (HYPOX) by the transauricular route at 35 days old and the experiments were done 6 weeks later. Saline (controls), dextrose (10%), 2-deoxy-D-glucose (2-DG, 500 mg per kg) and glucagon (0.9–1.0 U) were all given i.p. in aqueous medium (1 ml) 30 min before killing. Plasma corticosterone levels were low and near the lower limit of detection (5 nmol l⁻¹) for all ADRX and HYPOX experimental groups (see Table 1 for typical levels in intact animals). Plasma glucose was assayed by the hexokinase method ('Glucose Rapid Test', Roche). Values are mean ± s.e.m.

* $P < 0.05$ versus controls; † $P < 0.01$ versus controls.

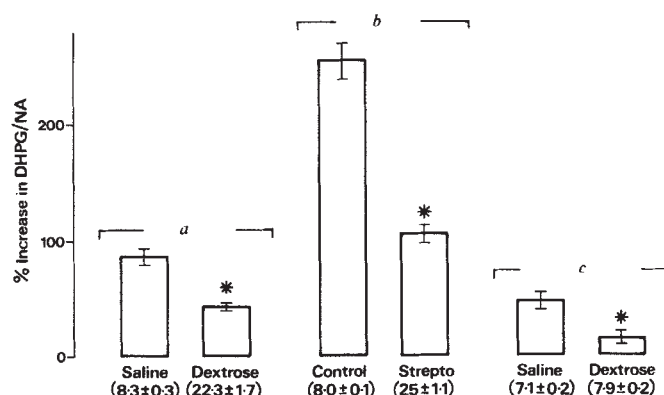


Fig. 2 The effect of elevated glucose levels (shown in parentheses) on the increase in hypothalamic DHPG/NA (expressed as % of control value): a, 15 min after cold-swim stress (0 °C, 2 min) in intact rats treated 20 min before killing with saline (1 ml i.p.) or dextrose 1 ml of 50% solution i.p.); b, 30 min after yohimbine treatment (8 mg per kg i.p.) in control and streptozotocin-induced diabetic rats; and c, in adrenalectomized rats 20 min after administration of saline (1 ml i.p.) or dextrose (1 ml of 10% solution i.p.). Means ± s.e.m. are shown. There were six animals per group (except for the adrenalectomy/saline group where $n = 10$).

* $P < 0.01$ compared with control.

The data presented here reveal significant interrelationships between hypothalamic NA neuronal activity and plasma glucose levels and we suggest that this may be the mechanism whereby the brain regulates its energy supply. We conclude that the brain glucose control system is mediated via central NA neuronal activity, which is stimulated by neuroglycopenia and suppressed by glucose excess.

Received 1 November; accepted 2 December 1983.

1. Brown, J. *Metabolism* **11**, 1098–1112 (1962).
2. Korec, R. in *Experimental Diabetes Mellitus in the Rat* (ed. Korec, R.) 83–87 (Slovak Academy of Sciences, Bratislava, 1967).
3. Benzo, C. A. *Life Sci.* **32**, 2509–2515 (1983).
4. Smythe, G. A., Bradshaw, J. E. & Vining, R. F. *Endocrinology* **113**, 1062–1071 (1983).
5. Smythe, G. A., Duncan, M. W., Bradshaw, J. E. & Cai, W. Y. *Endocrinology* **110**, 376–383 (1982).
6. Smythe, G. A., Bradshaw, J. E., Cai, W. Y. & Symons, R. G. *Endocrinology* **111**, 1181–1191 (1982).
7. Feldman, M., Kiser, R. S., Unger, R. H. & Li, C. H. *New Engl. J. Med.* **308**, 349–353 (1983).
8. Warsh, J. J., Li, P. P., Godse, D. D. & Chang, S. *Life Sci.* **29**, 1303–1307 (1981).
9. Anden, N. E., Pauksens, K. & Svensson, K. *J. Neural Transm.* **55**, 111–120 (1982).
10. Shimazu, T. *Diabetologia Suppl.* **20**, 343–356 (1981).

Monoclonal antibodies identify blastemal cells derived from dedifferentiating muscle in newt limb regeneration

Chris R. Kintner & Jeremy P. Brockes

MRC Cell Biophysics Unit, 26–29 Drury Lane, London WC2B 5RL, UK* and Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

Blastemal cells arise after the amputation of limbs or tails in urodele amphibians. These histologically undifferentiated mesenchymal cells divide and subsequently differentiate to regenerate a new appendage. Various studies (reviewed in ref. 1) indicate that blastemal cells arise from tissues near the site of amputation, including muscle, cartilage, nerve and dermis. The multinucleated myofibre, however, is a controversial source of blastemal cells². The suggestion that myofibres can dedifferentiate is based on their histological appearance during the early stages of limb regeneration^{3–5}. This is contrary to the widely accepted view of muscle regeneration in higher vertebrates which attributes it to satellite cells^{6,7}. One prediction of the dedifferentiation hypothesis is that a population with properties of both myofibres and blastema cells should be present during the early stages of regeneration. Here we describe the isolation of two monoclonal antibodies, one that recognizes an antigen found only in myofibres and another that recognizes an antigen restricted to blastemal cells. By using these antibodies as cell markers, we can detect a small population of cells in the regenerating limbs of adult newts that bear both the myofibre and blastemal cell antigens. The time and location of these double-labelled cells supports the idea that blastemal cells originate, in part, by dedifferentiation of myofibres.

Monoclonal antibodies were generated from mice immunized with mid-bud stage blastemas⁸ derived from adult newts (*Notophthalmus viridescens*) after amputation through the upper forelimb. The properties of one hybridoma, named 22/18, that recognizes an antigen restricted to blastemal cells, are summarized in Table 1. 22/18 does not bind to any tissue in the mature limb when assayed by indirect immunofluorescence on tissue sections. 22/18 staining is first detected about 4 days after amputation in cells of the nerve sheath near the plane of amputation. During the next 10 days, the number of cells labelled with 22/18 increases and includes the majority of the blastemal cells accumulating under the wound epidermis (Fig. 1a, b), as well as mononucleated cells among muscle and nerve at the base of

the blastema. At the late bud stage of regeneration⁸, only a minor population of blastemal cells is labelled with 22/18 and it is located between the epidermis and the central core of the blastema. As blastemal cells differentiate into mature tissue the 22/18 staining is lost. The staining patterns of this reagent will be reported in detail elsewhere.

A monoclonal antibody that recognizes an antigen restricted to myofibres was derived from mice immunized with homogenized muscle from adult newts. Table 1 gives the isolation and properties of this antibody, named 12/101. 12/101 reacts with myofibres (Fig. 1c, d) but not with any other tissue in amputated or normal limbs as assayed by indirect immunofluorescence in tissue sections. In some conditions of fixation the striations in myofibrils were clearly visible and when this occurred the 12/101 staining was also striated, suggesting that the antigen is associated with the contractile apparatus.

Longitudinal tissue sections were cut from regenerating limbs 14 days after amputation and stained with both 12/101 and 22/18 in conditions that allowed their separate detection on the same section with antibodies conjugated to rhodamine and fluorescein (see Fig. 1 legend). The myofibres in the mature tissue at the base of the limb were stained normally with 12/101. The 12/101 staining became fragmented and discontinuous, however, within the region between mature tissue and the blastema. Figure 1 shows an example of this region photographed under low (e–g) and high (h–j) magnification and oriented so that the mature tissue lies towards the top of the photograph and the blastema towards the bottom. Some of the 12/101 staining in this region was in cytoplasm associated with a single nucleus as detected by Hoescht dye (examples are designated with arrowheads in Fig. 1f and i). Of these cells, some also stain unambiguously with the blastemal cell marker 22/18 (Fig. 1g, j).

Thus, the region of the regenerating limb between the blastema and mature tissue contains apparently mononucleated cells that stain for both the myofibre and blastemal antigens. This region of the regenerating limb also contains intact myofibres that are associated with several nuclei and stain with 12/101 (examples are designated by arrows in Fig. 1f). In some cases, these myofibres are also stained with 22/18, usually at the end closest to the blastema (arrows in Fig. 1g). These double-labelled, multinucleated cells have the properties expected for myofibres that have initiated the process of dedifferentiation.

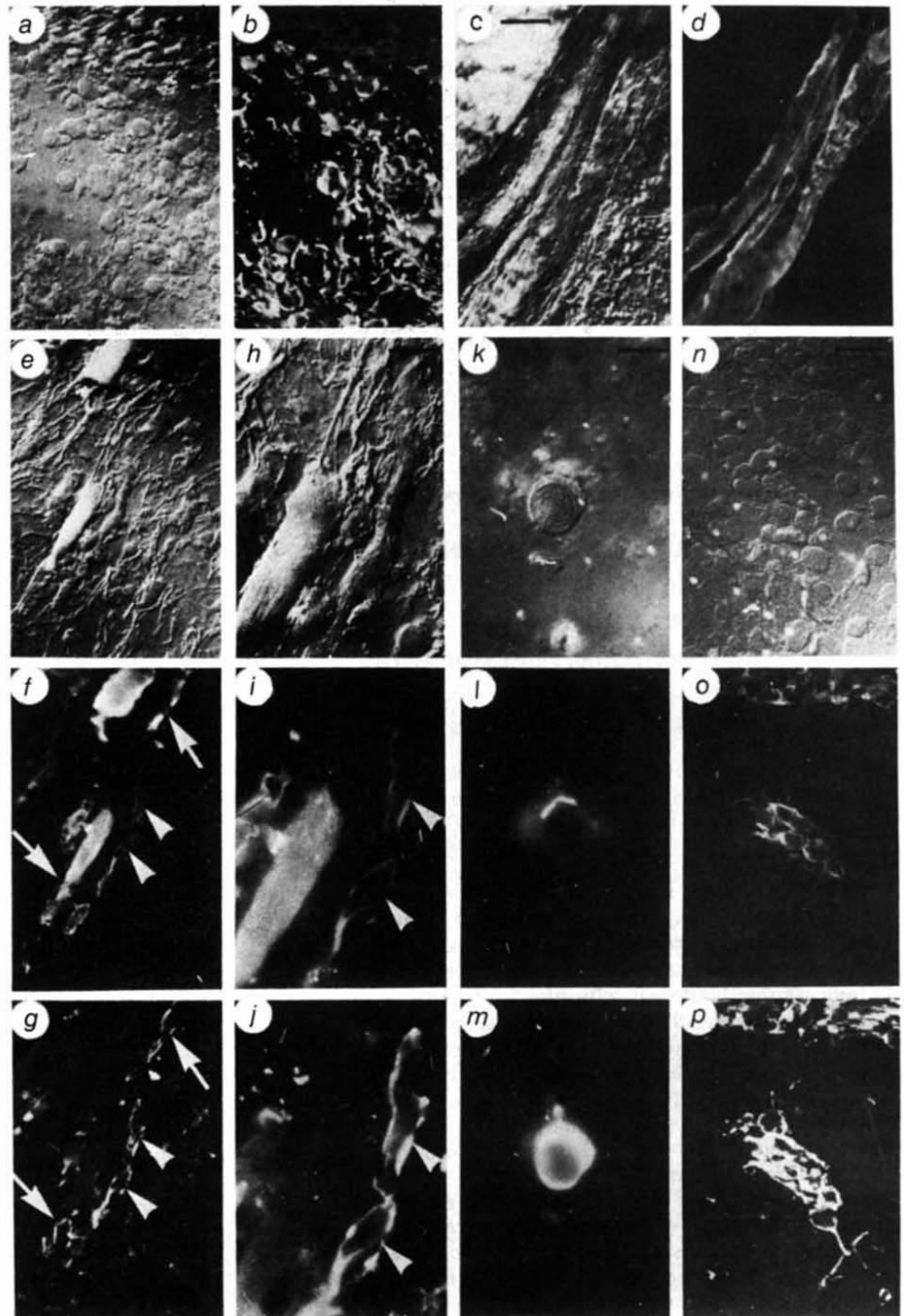
The double-labelled mononucleated cells appear by several criteria to be blastemal cells arising from myofibre dedifferentiation. First, the location of these cells is definitely restricted to the region of the regenerating limb between the blastema and intact myofibres. Double-labelled cells are not detected in the stump among intact myofibres, nor do they occur in the growing bud but histologically undifferentiated blastema where a majority of the cells stain with the blastemal cell marker.

Second, double-labelled cells appear at a specific time during regeneration. They are detectable in regenerating limbs at 12 and 14 days after amputation but not at 4 or 7 days, when wound healing takes place in the amputated limb. At this earlier stage, there are fragments of myofibres stained with 12/101 in proximity to mononucleated cells stained with 22/18 but the two populations never overlap. Not only do the double-labelled cells appear after wound healing but they appear before muscle differentiation occurs in the blastema (as discussed below). Thus, it is likely that the double-labelled cells detected in regenerating limbs 14 days after amputation are associated with dedifferentiation rather than with differentiation or wound healing.

Third, double-labelled cells are detected after enzymatic dissociation of tissue followed by antibody staining of isolated cells (Fig. 1k–m). These cells were reproducibly detected at a low frequency (3 double-labelled cells per 339 22/18-positive cells in the experiment of Fig. 1). It could otherwise have been argued that the double-labelled cells observed in sections were artefacts that arise when a cell appears to bear an antigen that belongs to a closely apposed process of another cell.

* Present address.

Fig. 1 **a**, Tissue section of an early-bud regenerate stained with 22/18. The area of the blastema under the wound epidermis (upper field) was photographed under Nomarski optics (scale bar, 50 μ m) or rhodamine optics (**b**) to show 22/18 staining. Note that blastema cells are stained while epidermis is not. **c**, Tissue section of a mature limb stained with 12/101. An area of the limb that includes nerve (left field), muscle (centre field) and dermis (right field) was photographed under Nomarski optics (scale bar, 50 μ m) or fluorescein optics (**d**) to show 12/101 staining. Note that only muscle stains with 12/101. **e**, Tissue section of a regenerating limb 14 days after amputation stained with 22/18 and 12/101. An area between the blastema and intact myofibrils was photographed with Nomarski optics (scale bar, 50 μ m), fluorescein optics (**f**) to show 12/101 staining and rhodamine optics (**g**) to show 22/18 staining. The arrowheads designate apparently mononucleated cells that stain with both 12/101 and 22/18. The arrows designate multinucleated myofibres that stain with both 12/101 and 22/18. Note that the muscle in the upper field stains with 12/101 but not 22/18 and the blastemal cells in the bottom field stain with 22/18 but not 12/101. **h**, **i**, **j**, The centre field of **e**, **f** and **g** photographed in the same conditions but at higher magnification (scale bar, 20 μ m). Note the two cells designated with arrowheads stain with both 12/101 and 22/18. **k**, A dissociated cell stained with both 12/101 and 22/18 and photographed under Nomarski optics (scale bar, 20 μ m), fluorescein optics (**l**) to detect 12/101 staining and rhodamine optics (**m**) to detect 22/18 staining. Note the fibrillar nature of 12/101 staining, which is observed in some double-labelled cells and is consistent with the ultrastructural description of some blastemal cells containing fragments of myofibrils⁵. **n**, Tissue section of a palette stage regenerate stained with both 22/18 and 12/101. An area of blastema near a region of muscle differentiation was photographed under Nomarski optics, fluorescein optics (**o**) to show 12/101 staining or rhodamine optics (**p**) to show 22/18 staining. At this stage of regeneration, 22/18 only stains a subpopulation of blastemal cells. Note the two regions where cells stain with both 22/18 and 12/101.



Methods: Staining of tissue sections and dissociated cells from regenerating limbs by indirect immunofluorescence with the 22/18 and 12/101 antibodies. Adult newts (*Notophthalmus viridescens*, from Lee's Newt Farm in Tennessee) were bilaterally amputated through the upper forelimbs and allowed to regenerate for the times specified. For tissue sections, the regenerating limbs were fixed for 45 min by immersion in 0.01 M Na periodate, 0.1 M lysine and 0.05% paraformaldehyde¹² in HEPES-buffered saline (HBS: 120 mM NaCl, 3 mM glucose, 2 mM KCl, 5 mM HEPES pH 7.2 and 10 mM CaCl₂), washed in HBS for 30 min and then frozen in embedding medium. Frozen sections (12 μ m) were cut with a cryostat, collected onto coverslips coated with Gatenby's solution, dried and stored at -70 °C. Dissociated cells were prepared from 20 animals, 17 days after bilateral amputation of the forelimbs. The portion of the regenerating limb lying 1-4 mm from the amputation site was removed, finely chopped with a razor blade and digested in 0.5 mg ml⁻¹ papain (a fraction obtained after ion exchange chromatography of papain from Calbiochem on CM-cellulose at pH 5.0 with a linear 0.1-1.0 M ammonium acetate gradient), 0.88 mg ml⁻¹ cysteine in HBS for 15 min at 28 °C. Cells released by this procedure were filtered through nylon mesh, washed once with HBS, collected by centrifugation, dried onto glass coverslips and fixed with acetone. For detection of double-labelled cells, tissue sections or dissociated cells were simultaneously stained with 22/18 and 12/101 using reagents specific for μ - or γ -chains of mouse immunoglobulin (see Table 1) such that 22/18 binding was detected with rhodamine and 12/101 with fluorescein. Tissue sections or cells were exposed to ascites fluid of 22/18 (1:200) and 12/101 (1:100), followed by rabbit anti-mouse IgM, μ -chain specific (1:200, Miles) and then to rhodamine-conjugated goat anti-rabbit IgG (1:150, Cappel) and fluorescein-conjugated goat anti-mouse IgG, γ -chain specific (1:75, Cappel). Reagents were diluted in phosphate-buffered saline with 10% fetal bovine serum, and all antibody incubations were for 45 min at room temperature. Staining with 22/18 or 12/101 alone was performed with the appropriate reagents as used for double staining. Controls in which the monoclonal antibodies were omitted, or control monoclonal antibodies were used, routinely gave no staining except occasionally of the epidermis. Stained sections or cells were postfixed in cold acid alcohol, stained with Hoescht dye 33258 (0.5 mg ml⁻¹) and mounted in 90% glycerol with phosphate-buffered saline. Stained material was viewed and photographed using a Zeiss photomicroscope equipped with epifluorescence, a filter system to separate fluorescein, rhodamine and Hoescht dye fluorescence, and Planapo $\times 63.2$ and Plan Neofluar $\times 25$ objectives with Nomarski facilities.

Table 1 Isolation and properties of monoclonal antibodies 22/18 and 12/101

Monoclonal antibody designation	Immunogen	Staining in normal limb	Additional specificity in regenerating limbs	Subclass
22/18	Crude membrane fraction of mechanically dissociated cells from mid-bud blastemas	None	Majority of blastemal cells at early stages. Minority of blastemal cells at later stages of regeneration	IgM
12/101	Extracts of muscle from adult newts	Myofibrils	Small population of cells at early and late stages of regeneration.	IgG1

Spleen cells from mice injected three times with immunogen were used to generate hybridomas by fusion to the sp2/0 myeloma line as described in ref. 11. Hybridomas showing a desired specificity by immunofluorescence on tissue sections were cloned several times by limiting dilution and injected into mice to obtain ascites fluid¹¹. Frozen sections of normal or amputated limbs were stained by indirect immunofluorescence with 22/18 and 12/101 as described in Fig. 1 legend. Subclass determination was by immunofluorescence using rabbit anti-mouse IgG1, IgG2A, IgG2B or IgM (Miles) as the second layer and rhodamine-conjugated goat-anti-rabbit immunoglobulin as the third layer.

Cells that stain with both 12/101 and 22/18 were also detected in the blastema during the later stages of regeneration, when differentiation occurs. An example of double-labelled cells that are present in sections of palette-stage⁸ regenerating limb is shown in Fig. 1*n-p*. The location of double-labelled cells in the blastema during later stages of regeneration is restricted to regions just distal to immature myofibres. On one hand, the immature myofibres just proximal to the double-labelled cells stain intensely with the muscle marker but not with the blastemal cell marker. On the other hand, the blastemal cells that stain with 22/18, but are located in areas of the blastema where differentiation has not occurred, do not stain with 12/101. Because the double-labelled cells are just distal to immature myofibres, they are likely to be blastemal cells that have initiated the early stages in myogenesis.

Thus, we have generated two monoclonal antibodies that serve as markers for myofibres and blastemal cells. A small population of apparently mononucleated cells that appear during early stages of regeneration in a restricted region of the regenerating limb are stained with both. We interpret these double-labelled cells to be blastemal cells arising by myofibre dedifferentiation. Although other interpretations may explain the occurrence of these cells, they seem much less likely. One possibility is that they are phagocytic cells that have taken up exogenous antigen. An association between the two markers is not found during stages of wound-healing when myofibre-positive debris and cells positive for 22/18 are present in close proximity in the wound area. Furthermore, the association between the myofibre and blastemal cell markers is specific for one blastemal cell type in the sense that double-labelled cells are not detected when sections are stained with 12/101 and for another marker restricted to a population of blastemal cells abundant in the area of dedifferentiation. This latter marker is defined by a different monoclonal antibody and is detected on muscle fibroblasts and a population of blastemal cells not stained with 22/18 (unpublished observations). The double-labelled cells are not likely to be the satellite cells that appear to mediate muscle regeneration in higher vertebrates. Studies on skeletal muscle in adult newts have failed to detect satellite cells, at least as defined by the criterion of mononucleated cells within the muscle basal lamina^{9,10}. Furthermore, it seems unlikely that satellite cells at this stage, however defined, would transiently express an antigen that appears to be associated with the contractile apparatus. Most important, the presence of double-labelled multinucleated cells is suggestive as an intermediate between myofibres and double-labelled blastoma cells.

Finally, it is interesting that after the onset of differentiation in the blastema, 22/18-positive blastemal cells differentiate back into muscle. It is possible that the 22/18-positive cells destined for myogenesis are exclusively the progeny of cells that arose by myofibre dedifferentiation (referred to as clonal modulation in ref. 1). This possibility cannot be tested using only the 22/18 antibody as a cell marker because other tissues such as nerve sheath cells appear to give rise to 22/18-positive blastemal cells. Nonetheless, the isolation and use of other monoclonal antibodies as cell markers may allow the description of such lineages.

We thank T. Stevens and S. Joslin for technical assistance, Drs A. J. Hudspeth and R. Lewis for communicating unpublished results on papain fractionation for amphibian tissue dissociation and C. E. Maier for instruction in the surgery for newt regeneration. Research at Caltech was supported by grants from the NIH and the Pew Foundation.

Received 7 September; accepted 5 December 1983.

- Wallace, H. *Vertebrate Limb Regeneration* (Wiley, Chichester, 1981).
- Hay, E. D. in *Regeneration of Striated Muscle and Myogenesis* (eds Mauro, A., Shafiq, S. & Milhorat, A. T.) 3-24 (Excerpta Medica, Amsterdam, 1970).
- Thornton, C. S. *J. Morph.* **62**, 17 (1938).
- Hay, E. D. *Dev. Biol.* **1**, 555 (1959).
- Lentz, T. L. *Am. J. Anat.* **124**, 447 (1969).
- Bischoff, R. in *Muscle Regeneration* (ed. Mauro, A.) 13-29 (Raven, New York, 1979).
- Konigsberg, U. R., Lipton, B. H. & Konigsberg, I. R. *Dev. Biol.* **45**, 260 (1975).
- Iten, L. E. & Byrant, S. V. *Wilhelm Roux Arch. Entw. Mech. Org.* **173**, 263 (1973).
- Hay, E. D. & Doyle, C. M. *Anat. Rec.* **175**, 339 (1973).
- Popiela, H. J. *exp. Zool.* **198**, 57 (1976).
- Moore, H.-P. H., Fritz, L. C., Raftery, M. A. & Brockes, J. P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1673 (1982).
- McLean, I. W. & Nakane, P. K. *J. Histochem. Cytochem.* **22**, 1077 (1974).

Transforming *ras* genes from human melanoma: a manifestation of tumour heterogeneity?

Anthony P. Albino, Renee Le Strange, Allen I. Oliff, Mark E. Furth & Lloyd J. Old

Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA

Variability in the phenotype of cells comprising individual tumours is a striking feature of animal and human cancer and is generally referred to as tumour heterogeneity¹. Studies of clonally derived cell populations from tumours that originated presumably from a single transformed cell have shown that tumours are made up of cells that differ in a variety of traits, including drug resistance, antigen expression and metastatic potential²⁻⁶. The origin and maintenance of tumour heterogeneity are unclear, but mutational and epigenetic mechanisms are thought to be involved. Here we report the results of a search for transforming genes in human melanoma which have raised the possibility that *ras* gene activation follows the same variable pattern as other traits involved in tumour heterogeneity. DNA from 4 of 30 melanoma cell lines yielded transforming *ras* genes in the NIH/3T3 assay. Of five cell lines originating from separate metastatic deposits of a single patient, only one contained activated *ras*, indicating heterogeneity in *ras* activation in this case and suggesting that *ras* activation was not involved in tumour initiation or maintenance in this patient.

Table 1 summarizes the transfection experiments with high molecular weight (M_r) DNA extracted from cultured melanoma cells. The efficiency of primary transfection from the positive

cell lines (SK-MEL-93, -119, -146 and -147) ranged from 0.01 to 0.10 focus-forming units (FFU) per microgram of DNA ($\text{FFU } \mu\text{g}^{-1}$). The efficiency of focus formation was not significantly altered in secondary and tertiary rounds of transfection. The transformed NIH/3T3 cells produced rapidly growing solid tumours upon injection into NFS/Swiss weanling mice. Table 1 also summarizes the transfection experiments with DNA extracted from five SK-MEL-93 cell lines (DX-1, -2, -3, -4 and -6) derived from separate metastatic foci of patient DX, and with DNA from DX's normal skin fibroblasts and Epstein-Barr virus EBV-transformed peripheral blood lymphocytes. In a series of 11 separate transfection experiments, DNA from only one of these lines, DX-3, transformed NIH/3T3 cells.

Figure 1 shows DNA blots of NIH/3T3 cells after three cycles of transfection with oncogenes from SK-MEL-146 and SK-MEL-93, probed for human-derived repetitive (*Alu*-type) DNA sequences⁷. The patterns obtained after digestion with five restriction endonucleases revealed that SK-MEL-146 and SK-MEL-93 tertiary transfectants had no bands in common, indicating that the two parental melanoma lines contributed different oncogenes. Independent isolates of secondary and tertiary transfectants with the SK-MEL-93 oncogene also showed the 1.9 kilobase (kb), 6.0 kb and 9.1 kb *EcoRI**Alu*-containing human DNA fragments seen in Fig. 1 (lane 2). Similarly, independent isolates of secondary and tertiary SK-MEL-146 transfectants conserved 10 kb and 23.5 kb *EcoRI**Alu*-containing fragments (lane 1).

Figure 2 shows the results of probing DNA from tertiary transfectants containing SK-MEL-146 and SK-MEL-93 oncogenes, as well as DNA from the two parental melanoma lines, for *ras* sequences. In the experiment shown in Fig. 2a, *Bam*HI-digested DNAs were hybridized with a probe derived from the *v-ras*^H gene of Harvey murine sarcoma virus (MuSV) (BS9, ref. 8). The endogenous *c-ras*^H genes of SK-MEL-146 and -93 are visualized in lanes 1 and 3, respectively. A human-derived fragment containing *ras*^H-related sequences is present in DNA obtained from tertiary transfectants containing the SK-MEL-146 oncogene (lane 2), but not in transfectants containing oncogenes from SK-MEL-93 (lane 4), SK-MEL-119 or SK-MEL-146 (data not shown). Since the *ras*^H-related sequences migrated as an 8.5 kb *Bam*HI fragment in SK-MEL-146, but as an 11.0 kb *Bam*HI fragment in tertiary SK-MEL-146 transfectants (see lane 2), we assume that sequences flanking the *c-ras*^H oncogene were rearranged during the transfection process. Figure 2b is a replicate of the DNA blot shown in Fig. 2a, hybridized with a probe specific for the human *ras*^N genes⁹. The endogenous *ras*^N-related sequences appear as a 7.8-kb *Bam*HI fragment in SK-MEL-146 (lane 5) and SK-MEL-93 (lane 7), as well as in SK-MEL-119 and -147 (data not shown). A transfer of human *ras*^N-related sequences has occurred in tertiary SK-MEL-93 transfectants (lane 8), but not in tertiary SK-MEL-146 transfectants (lane 6). *ras*^N-related sequences were also found in secondary transfectants of SK-MEL-119 and -147 (data not shown). Replicate blots of those shown in Fig. 2, hybridized with *ras*^K sequences from Kirsten MuSV (probe HiHi 3) (ref. 10) failed to detect transfer of human *ras*^K sequences to NIH/3T3 transfectants.

Figure 3 shows that NIH/3T3 cells transformed by *ras*^H from SK-MEL-146 cells synthesize a novel, rapidly migrating protein p21-*ras* (arrow b, lanes 5, 6) detected by immunoprecipitation with two p21 monoclonal antibodies¹¹. Nontransformed 3T3 cells contain little Harvey-type p21-*ras* protein (lane 2), but significant amounts of Kirsten-type p21-*ras* (arrow a), as detected by a broadly cross-reacting *ras* monoclonal antibody (lane 3). An altered p21-*ras* protein has also been detected in transfectants carrying the *ras*^N oncogene from SK-MEL-93 (data not shown). Presumably, the novel p21-*ras* species result from structural mutations. As has been suggested in other cases, it is likely that qualitative changes in p21-*ras* function account for the transforming potential^{12,13}.

The present study extends the list of human tumour types that yield transforming DNA to malignant melanoma. As in

Fig. 1 Comparison of human *Alu*-containing DNA fragments in mouse tertiary transfectants of SK-MEL-146 and SK-MEL-93. Lanes 1, 3, 5, 7 and 9 show DNA from tertiary SK-MEL-146 transfectants; lanes 2, 4, 6, 8 and 10 show DNA from tertiary SK-MEL-93 transfectants. The size markers indicated are *Hind*III fragments of λ phage DNA.

Methods: Twelve μg samples of the DNAs were digested with the indicated endonuclease, fractionated by electrophoresis through 0.75% agarose gels at 20 °C for 16 h, and transferred to nitrocellulose paper²⁷. The filters were incubated with a nick-translated²⁸ ³²P-labelled BLUR 8 *Alu*

probe in the following conditions: 50% formamide, 5 $\times$ SSC, 1 $\times$ Denhardt's, 25 $\mu\text{g ml}^{-1}$ salmon sperm DNA, 10% dextran sulphate for 16 h at 40 °C. Blots were washed at room temperature for 1 h with three changes of 2 $\times$ SSC, 0.1% SDS, then at 50 °C for 1 h with two changes of 0.1 $\times$ SSC, 0.1% SDS. The blots dried and exposed to Kodak XAR-5 film with Dupont Lightning-Plus intensifying screen at -70 °C.

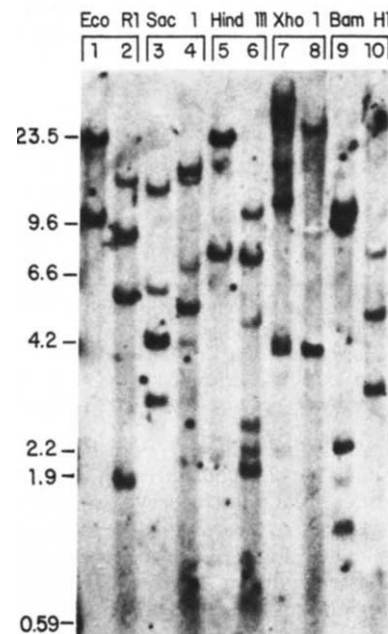


Table 1 Transfection of melanoma DNA

Cell type*	Efficiency of primary transfection (FFU μg^{-1})†	Tumorigenicity of transfectants‡
SK-MEL-119	0.04, 0.068, 0.1	5/5
SK-MEL-146	0.04, 0.072, 0.1	8/8
SK-MEL-147	0.01, 0.012	3/3
SK-MEL-23, 28, 30, 37, 44, 47, 64, 65, 75, 100, 113, 118, 122, 129, 131, 143, 153, 164, 166, 167, 169, MeWo	<0.003 (in the case of each cell line)	0/30
SK-MEL-93 (DX-3)	0.01, 0.04	8/8
(DX-1)	<0.003	0/3
(DX-2)	<0.003	0/3
(DX-4)	<0.003, <0.003	0/6
(DX-6)	<0.003, <0.003, <0.003	0/6
DX skin fibroblast	<0.003	0/3
DX EBV-transformed B cells	<0.003	0/3

* Melanoma cell lines were derived and maintained as described previously⁶.

† 400 μg or high molecular weight DNA (>50 kb) extracted from cells as described by Lowy²³ was transfected to NIH/3T3 cells by the calcium phosphate technique²⁴ as follows: 10 μg precipitated human DNA was applied to each of 30 35-mm Petri dishes containing 2.5×10^5 NIH/3T3 cells: the cultures were incubated at 37 °C for 6 h with a subsequent 30-min shock in 10% dimethyl sulphoxide²⁵, or for 20 h without shock. After incubation, the cells were trypsinized and replated in 100-mm Petri dishes and fed every 3 days. After 2–3 weeks, foci of transformed NIH/3T3 cells were isolated with cloning cylinders, grown to mass culture and used as source of DNA for further analysis. DNA competency in transfection was confirmed by transfer of human thymidine kinase (TK) gene to LMTK⁻ cells with a portion of the same DNA precipitate used in transfection assays. Selection of TK⁺ colonies in 1% hypoxanthine-aminopterin-thymidine HAT medium showed the transfer efficiency of human TK gene to be 0.1–0.5 colonies per μg of DNA²⁶.

‡ Tumorigenicity of transformed cells was determined by injecting $1-2 \times 10^6$ cells (0.1–0.2 ml) subcutaneously into 4–6-week-old NFS mice. Tumours were detected after 1–2 weeks. In the case of cultures scored negative for transformed foci, the cells were trypsinized, grown to mass culture and injected into weanling NFS mice (5×10^6 cells per mouse) to determine the presence of any transformed cells not identified microscopically. Animals were observed for 1–2 months.

Fig. 2 Detection by Southern blot analysis of *ras* gene sequences in DNA from donor melanomas and from tertiary transfectants. Lanes 1, 5, SK-MEL-146 DNA; lanes 2, 6, tertiary SK-MEL-146 transfectant DNA; lanes 3, 7, SK-MEL-93 DNA; lanes 4, 8, tertiary SK-MEL-93 transfectant DNA. The 3.0-kb bands visualized in lanes 2 and 4 correspond to the endogenous murine *c-ras^H* gene. This band is also visualized in control NIH/3T3 cells (data not shown). The size markers indicated are *Hind*III fragments of λ phage DNA. Twelve μ g DNA samples were digested with *Bam*HI and electrophoresed, blotted and washed as described in Fig. 1. The *a* filter was hybridized with a nick-translated 32 P-labelled *ras^H* specific probe (BS9); the *b* filter was hybridized with a nick-translated 32 P-labelled *ras^N* specific probe (designated R)⁹.

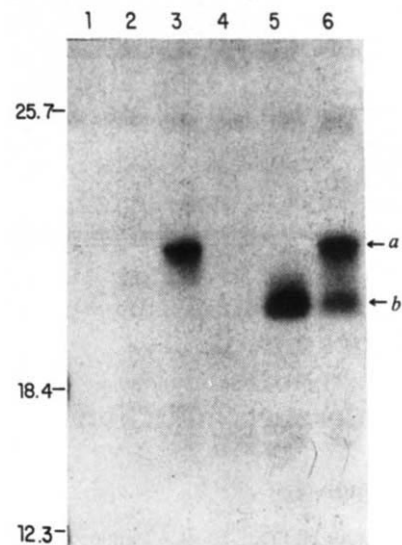
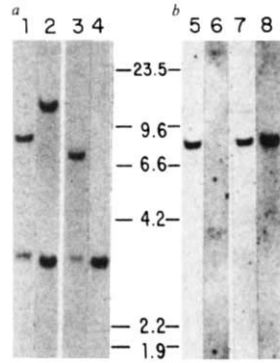


Fig. 3 Expression of p21 protein in NIH/3T3 cells transformed by human melanoma DNA. Lane 1, NIH/3T3 cells plus normal rat IgG; lane 2, NIH/3T3 cells plus Y13-238 monoclonal antibody; lane 3, NIH/3T3 cells plus Y13-259 monoclonal antibody; lane 4, tertiary transfectant of SK-MEL-146 plus normal rat IgG; lane 5, tertiary transfectant of SK-MEL-146 plus Y13-238 monoclonal antibody; lane 6, tertiary transfectant of SK-MEL-146 plus Y13-259 monoclonal antibody. Arrows: *a*, major p21-*ras* species of NIH/3T3 cells; *b*, novel p21-*ras* species encoded by activated oncogene in NIH/3T3 tertiary transfectant from SK-MEL-146. Molecular weight markers ($\times 10^{-3}$) (14 C-labelled polypeptides from Bethesda Research Laboratories).

Methods: Cells were labelled by incubation for 18 h at 37°C in methionine-free medium containing 2% dialysed fetal calf serum and 100 μ Ci ml⁻¹ 35 S-methionine. Lysates were prepared and immunoprecipitation was carried out essentially as described by Furth *et al.*¹¹. Portions of lysates containing equal amounts of incorporated isotope (10×10^6 c.p.m. per reaction) were incubated with 0.5 μ g normal rat immunoglobulin or purified monoclonal rat IgG Y13-238 or Y13-259¹¹ for 18 h at 4°C. One monoclonal antibody, Y13-238 (ref. 11), precipitates the product of Harvey-type *ras*, but does not precipitate the p21-*ras* encoded by the *ras^K* genes of Kirsten MuSV or of normal rodent cells. A second monoclonal antibody, Y13-259, precipitates the products of *ras^H*, *ras^K* and *ras^N* (refs 9, 11). Antigen-antibody complexes were collected by precipitation with formalin-fixed *Staphylococcus aureus* bacteria precoated with rabbit antibodies against rat IgG (ref. 11). After extensive washing, the bacteria were boiled for 3 min in sample buffer containing 2% SDS and 2-mercaptoethanol. Samples were subjected to electrophoresis on a 15% (w/v) polyacrylamide gel containing 0.1% SDS by the method of Laemmli *et al.*³⁰. Radiolabelled polypeptides were revealed by fluorography. Dried gels were exposed to Kodak XAR-5 film at -70°C using a Dupont Cronex screen.

Received 8 September 1983; accepted 6 January 1984.

- Heppner, G. H. & Miller, B. E. *Cancer Metastasis Rev.* **2**, 5-23 (1983).
- Nowell, P. C. *Science* **194**, 23-28 (1976).
- Fidler, I. J. & Hart, I. R. *Science* **217**, 998-1003 (1982).
- Shapiro, J. R., Yung, W. A. & Shapiro, W. R. *Cancer Res.* **41**, 2349-2359 (1981).
- Talmadge, J. E. *Cancer Metastasis Rev.* **2**, 25-40 (1983).
- Albino, A. P., Lloyd, K. O., Houghton, A. N., Oettgen, H. F. & Old, L. J. *J. exp. Med.* **154**, 1764-1778 (1981).
- Jelinek, W. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1398-1402 (1980).
- Ellis, R. W. *et al. J. Virol.* **36**, 408-420 (1980).
- Shimizu, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 2112-2116 (1983).
- Ellis, R. W. *et al. Nature* **292**, 506-511 (1981).
- Furth, M. E., Davis, J. L., Fleurdelys, B. & Scolnick, E. M. *J. Virol.* **43**, 294-304 (1982).
- Tabin, C. J. *et al. Nature* **300**, 143-148 (1982).
- Der, C. J. & Cooper, G. M. *Cell* **32**, 201-208 (1983).
- Der, C. J., Kontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637-3640 (1982).
- Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474-478 (1982).
- Santos, E., Tronick, S. R., Aaronson, S. A., Pulciani, S. & Barbacid, M. *Nature* **298**, 343-347 (1982).
- McCoy, M. S. *et al. Nature* **302**, 79-81 (1983).
- Shimizu, K., Goldfarb, M., Peruchio, M. & Wigler, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 383-387 (1983).
- Pulciani, S. *et al. Nature* **300**, 539-542 (1982).
- Yuasa, Y. *et al. Nature* **303**, 775-779 (1983).
- Cooper, G. M. *Science* **217**, 801-805 (1982).

previous studies with a range of other human tumours, the transforming genes from melanoma are members of the *ras* family¹⁴⁻¹⁹. Two types of *ras* genes were isolated from 4 of the 30 melanoma lines we tested, one *ras^H* allele and three *ras^N* alleles. Two types of *ras* oncogenes have also been isolated from human lung cancers, but in the case of lung cancer, *ras^K* (in addition to *ras^H*) rather than *ras^N* was the other *ras* allele isolated^{14,20}.

The frequency of *ras* gene activation in melanoma is 13%, a value not substantially different from the 10-20% frequency observed by other investigators with other human tumours (see refs 21, 22). Does this mean that *ras* is involved in the pathogenesis of only a subset of human cancers or could *ras* have a more general role in human cancer requiring approaches other than the NIH/3T3 assay to reveal? Another way to interpret the variability in *ras* activation in human cancer is that transforming *ras* genes are not involved in the origin and maintenance of human cancer, but arise as a consequence of the genetic instability of cancer cells and represent a manifestation of the high mutational rates of cancer cells¹⁻³. The finding that only one of five cell lines originating from the same patient had activated *ras* allele is interesting in this regard, but is also open to several interpretations: (1) The transforming *ras* gene was acquired during tissue culture passage, an objection that can be directed at all work on *ras* genes isolated from cell lines, but one that can be addressed by the study of noncultured melanomas. (2) The patient had two primary tumours, only one of which involved *ras* activation. A marker chromosome shared by the five lines would be important evidence for clonal origin of this patient's tumour, although the extensive chromosomal rearrangement and aneuploidy of melanoma cells hamper the search for a cytogenetic marker. Melanoma cells from this patient are known to express a clonotypic antigen not found on the patient's normal fibroblasts or B cells, nor on any of the 45 allogeneic melanoma cell lines tested⁶. Four of the five lines, including the line yielding the activated *ras^N* allele, expressed this clonotypic antigen, providing evidence for clonal derivation. (3) *ras* activation occurred relatively late in the course of tumour progression and was not involved in tumour initiation, maintenance or metastasis in this patient. According to the latter view, *ras* activation is a manifestation of tumour heterogeneity: a result rather than a cause of transformation. Studies on additional examples of multiple metastases from individual patients and on multiple clonal isolates from primary tumours will be necessary to pursue this idea of *ras* heterogeneity.

We thank Dan Littman (BLUR 8), Ronald Ellis (BS9, HiHi-3) and Michael Wigler (pNP-1) for supplying probes. Willa-jeanne McLean, Alison White, Tom Aldrich and Jane Becker provided technical assistance. This work was supported in part by NIH CA-08748, the Robert J. Kleberg and Helen C. Kleberg Foundation, American Cancer Society (MV-170) and Richard Molin Foundation for Cancer Inc.

22. Land, H., Parada, L. F. & Weinberg, R. A. *Science* **222**, 771-778 (1983).
23. Lowy, D. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5539-5543 (1978).
24. Loyer, A., Scangos, G. A. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 422-426 (1982).
25. Miller, C. L. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3346-3350 (1978).
26. Perucho, M. *et al. Cell* **27**, 467-476 (1981).
27. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
28. Maniatis, T., Jeffrey, A. & Kleid, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184-1188 (1975).
29. Rubin, C. M., Houck, C. M., Deininger, P. L., Friedmann, T. & Schmid, C. W. *Nature* **284**, 372-374 (1980).
30. Laemmli, U. K. *Nature* **227**, 680-685 (1970).

Essential requirement for major histocompatibility complex recognition in T-cell tolerance induction

Jonathan R. Lamb

Imperial Cancer Research Fund, Human Tumour Immunology Group, University College Hospital, Faculty of Clinical Sciences, University Street, London WC1E 6JJ, UK

Marc Feldmann

Imperial Cancer Research Fund, Tumour Immunology Unit, Department of Zoology, University College London, Gower Street, London WC1E 6BT, UK

The induction of T-cell responses involves the recognition of extrinsic antigen in association with antigens of the major histocompatibility complex (MHC), in mice and man, with different T cells recognizing antigen in association with either class I (H-2K/D, HLA-A, B, C) or class II (Ia, HLA-D/DR) MHC antigens¹⁻⁵. However, the requirement of MHC recognition in the induction of immunological tolerance remains ill defined. With human T helper clones recognizing synthetic peptides of influenza haemagglutinin (HA-1), we have investigated the nature of antigen-induced stimulation⁶, and antigen-induced antigen-specific unresponsiveness, immunological tolerance⁷. Tolerance is not due to cell death, as the cells remain responsive to interleukin-2 and is associated with the loss of T3 antigen from the cell surface⁸. Using monoclonal antibodies to the non-polymorphic regions of human class II antigens to inhibit the induction of T-cell tolerance we report here that induction of tolerance requires the recognition of MHC antigens.

Since T lymphocyte clones can be maintained in long-term culture it is now possible to analyse their mechanisms of antigen recognition and regulation in more detail than previously. We have produced a number of human influenza-specific T-cell clones with a diversity of function, such as help⁹, suppression¹⁰ and lymphokine production¹¹. They have been analysed for the fine specificity of their response to antigen. Clones HA1.4 and HA1.7 used in this study respond to a 24 amino acid peptide derived from the C-terminal region of HA-1 (residues 306-329) termed peptide 20, or p20† (ref. 12). This is an immunodominant region, as many clones raised with the intact haemagglutinin molecule react to this region¹³.

Depending on the nature and concentration of specific antigen, duration of culture and the presence of antigen-presenting cells, these clones can be either induced to proliferate, or can be rendered unresponsive⁷. Hence these clones are suitable for investigating the nature of the 'signals' necessary for T-cell activation or tolerance induction. Since recognition of MHC is of paramount importance for activation of T cells, we wanted to determine whether or not this was also the case for the induction of tolerance.

The human *HLA-D* region, corresponding to the mouse *H-2I* region, controls a number of different products, which have been described serologically under a variety of names. There are three *HLA-D* groups (reviewed in refs 2, 14, 15).

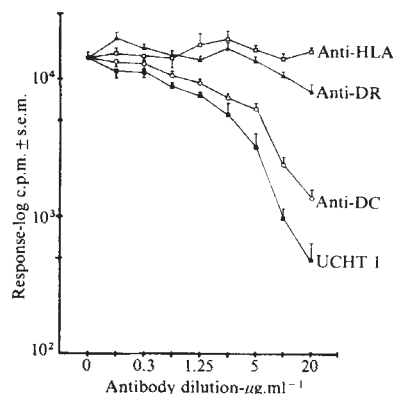


Fig. 1 Inhibition of antigen-induced proliferative response of cloned helper T cells reactive with HA peptide 20 by anti-HLA class II antibodies. Cells of HA1.7 (2.5×10^4 per ml) were stimulated with p20 ($1.0 \mu\text{g ml}^{-1}$) in the presence of irradiated E⁻ cells (2.5×10^4 per ml) in 96-well round-bottom culture plates. Antibodies (2A1, $\square$; DA-2, $\blacktriangle$; UCHT1, $\blacksquare$; and SG465, $\circ$) were added at the start of the cultures in the concentrations as indicated (0.15 – $20 \mu\text{g ml}^{-1}$ in doubling dilution) and left in for the duration of the culture. Proliferation was assayed as described in the legend to Table 1. The background responses of cloned T cells in the absence of irradiated E⁻ cells was <150 c.p.m. as was that of E⁻ cells alone cultured with antigen.

HLA-DR and *HLA-SB* correspond to the mouse *I-E* region, as judged by sequence homology and serological cross-reactivity. *HLA-DC/DS* (also known as *MT* and *MB, LB*), corresponds to the murine *I-A* subregion. There is evidence that *HLA-DR* (ref. 16), *SB* (ref. 17) and *DC* (D. D. Eckels, personal communication) all act as MHC targets of T-cell recognition by various clones.

The requirement for class II MHC determinants in T-cell activation can be demonstrated either by the addition of antigen-presenting cells bearing the appropriate MHC specificities or by the inhibition of T-cell proliferation by anti-class II antibodies specific for the various subregions¹⁸. The former approach cannot be used as tolerance is more readily induced in the absence of antigen-presenting cells⁷, and so serological inhibition, using rabbit anti-human Ia, and mouse monoclonal antibodies directed against class II antigens, either *HLA-DC/DS* or *HLA-DR*, was used. As controls, monoclonal anti-*HLA-A, B, C*, anti- β_2 -microglobulin, anti-T4 and anti-T3 antibodies were used.

The proliferative response of clone HA1.7 is inhibited by anti-class II, especially anti-*DC/DS* and thus involves recognition of *HLA-DC/DS* (Fig. 1). Anti-*HLA-A, B, C* had no effect, anti-*HLA-DR* had a minimal effect. This is presumably because there is a degree of cross-reactivity between the non-polymorphic components of *HLA-DR* and *HLA-DC*, as has been previously demonstrated biochemically^{19,20}.

One possible interpretation of the inhibition mediated by anti-*HLA-DC* was that it was due to a direct effect on T cells in the absence of antigen. This was excluded by experiments in which T cells were pretreated with antibodies and then washed before an antigenic challenge (Table 1). In these conditions, only anti-T3 inhibited T-cell proliferation. No enhancement or inhibition of proliferation was detected with either monoclonal or rabbit anti-class II antibodies. Therefore it would seem that the anti-class II antibodies mediated their effects by inhibiting antigen-induced activation¹⁸, and not by direct effect on the T cells.

Since human T-cell clones express class II antigens^{11,21,22} it was possible to investigate whether anti-class II antibodies could inhibit the induction of tolerance. Pretreatment of the clones with anti-*HLA-DC/DS* and rabbit anti-Ia blocked tolerance induction completely, whereas anti-*HLA-DR* had only a marginal effect (Table 2). Anti-MHC class I had no effect. These results indicate that tolerance involves recognition of MHC class

Table 1 Effect of antibodies on the induction of unresponsiveness in the absence of specific antigen

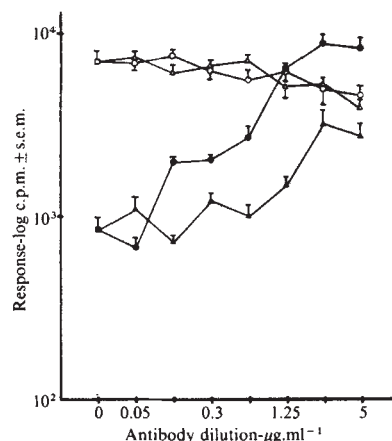
Antibody	Antigen	Response of cloned T helper cells, (c.p.m. $\pm$ s.e.m.)	
		E ⁻ + peptide 20	TCGF
—	—	11,454 $\pm$ 856	3,751 $\pm$ 497
—	p20	506 $\pm$ 104	5,677 $\pm$ 566
—	p11	10,924 $\pm$ 1,103	3,356 $\pm$ 331
Anti-HLA (2A1)	—	8,881 $\pm$ 804	4,309 $\pm$ 499
Anti-DC (SG465)	—	12,453 $\pm$ 1,862	2,264 $\pm$ 226
Anti-DR (DA2)	—	8,790 $\pm$ 329	2,744 $\pm$ 340
Rala	—	8,478 $\pm$ 1,123	2,994 $\pm$ 445
Anti-T3 (UCHT1)	—	698 $\pm$ 138	5,279 $\pm$ 267
Anti- β_2 M (M8)	—	11,926 $\pm$ 634	4,018 $\pm$ 482
Anti-T4 (MT321)	—	7,408 $\pm$ 612	3,406 $\pm$ 675

T lymphocyte clones reactive with the synthetic peptides of influenza virus haemagglutinin (HA) were isolated as described previously⁶. Briefly, peripheral blood mononuclear leukocytes (PBL) were stimulated for 6 days with HA (0.1 μ g ml⁻¹). Following enrichment on a discontinuous Percoll gradient (Pharmacia) the lymphoblasts were resuspended in RPMI-1640 (Gibco) supplemented with 10% pooled A⁺ serum and 20% T-cell growth factor (TCGF) and cloned by limiting dilution. Lymphoblasts were plated one cell every third well in sterile 60-well Microtest II trays (Falcon) with 10⁴ irradiated (2,500 rad) autologous (PBL) and 0.1 μ g ml⁻¹ of HA. At 7 days, growing clones were transferred to 96-well microtitre trays and subsequently to 24-well plates. At each transfer the clones received fresh TCGF every 3–4 days and irradiated histocompatible PBL and influenza virus (A/Texas/1/77) every 7 days. TCGF was prepared from tonsillar lymphocytes (1 $\times$ 10⁶ ml⁻¹) stimulated with 0.1% purified phytohaemagglutinin (Difco) in complete culture medium supplemented with 2.5% pooled A⁺ serum. Before use in experiments the clones were rested for 7 days after the addition of filler cells. To induce tolerance, cloned T cells were cultured at 10⁵ per ml in round bottom 96-well plates for 16 h with HA peptides (50 μ g ml⁻¹) or monoclonal antibodies (5 μ g ml⁻¹) for 16 h. Intact virus or haemagglutinin at high concentrations are toxic in culture and do not induce tolerance. Rabbit anti-Ia was used at a 1:50 dilution. After washing the cloned T cells (2.5 $\times$ 10⁴ per ml) were cultured with irradiated histocompatible sheep erythrocyte rosette negative (E⁻) lymphocytes (2.5 $\times$ 10⁴ per ml) that had been pulsed overnight with 1.0 μ g ml⁻¹ of p20. Following 60 h incubation the cultures were pulsed for 8–16 h with 1.0 μ Ci of ³H-thymidine (³HTdR) (Amersham) and collected onto glass fibre filters. Proliferation as correlated with ³H-TdR incorporation was measured by liquid scintillation spectroscopy. The results are expressed as mean counts per min (c.p.m.) $\pm$ s.e. of the mean of triplicate cultures. UCHT1 is specific for the T3 antigen present on all mature functional T cells³¹. 2A1 is against a nonpolymorphic determinant of HLA class I³². M8 is against β_2 -microglobulin³³. MT321 is against the helper/inducer subset (P. Rieber, personal communication), and rabbit anti-Ia is broadly reactive with human Ia-like antigens³⁴. DA-2 is reactive with a monomorphic determinant of HLA-DR (ref. 23) and SG465 is an antibody reactive with HLA-DC determinants²⁷.

Table 2 Anti-Ia antibody inhibits tolerance induction

Antibody	Antigen	Response of cloned T helper cells (c.p.m. $\pm$ s.e.m.)	
		E ⁻ + Peptide 20	TCGF
—	p20	674 $\pm$ 90	5,730 $\pm$ 437
—	p11	9,210 $\pm$ 524	4,461 $\pm$ 428
Anti-HLA (2A1)	p20	456 $\pm$ 82	4,150 $\pm$ 773
Anti-DC (SG 465)	p20	8,542 $\pm$ 267	2,368 $\pm$ 409
Anti-DR (DA2)	p20	2,225 $\pm$ 226	2,712 $\pm$ 276
Rala	p20	9,700 $\pm$ 972	2,697 $\pm$ 160
Anti-T3 (UCHT1)	p20	728 $\pm$ 137	5,141 $\pm$ 1016
Anti- β_2 M (M8)	p20	964 $\pm$ 258	3,589 $\pm$ 773
Anti-T4 (MT 321)	p20	880 $\pm$ 123	4,547 $\pm$ 71
—	—	10,591 $\pm$ 785	3,966 $\pm$ 486

Cloned T cells were preincubated with either peptide or antibody prior to assay or antigen-pulsed E⁻ cells as described in Table 1 legend. Antibody was added 6 h prior to the addition of antigen, and left in for the 16 h of antigen exposure before washing the cells three times. Proliferation was determined as described in the legend to Table 1.

**Fig. 2** Dose-response curve of the inhibition of tolerance induction by anti-HLA-DC and anti-HLA-DR antibodies. Cells of HA1.7 were pretreated with various concentrations of antibody (anti-HLA-DC, SG 465²⁷ ●, ○; anti-HLA-DR, DA2²³ ▲, △) and the antigen specific (●, ▲) and TCGF (○, △) responses assayed as described in the legend to Table 1.

II antigens. To determine more precisely which MHC class II antigens were involved, dose-response curves of anti-HLA-DC/DS and anti-HLA-DR antibodies on tolerance induction are compared. This was necessary as immunoprecipitation analysis has shown that the anti-DR monoclonal DA2 cross reacts weakly with DC/DS molecules^{19,20}, which would explain its partial effects on the induction of proliferation and tolerance. However, at lower concentrations (1.25 μ g ml⁻¹), only the anti-HLA-DC/DS monoclonal inhibits tolerance induction (Fig. 2), and thus the type of class II molecule recognized is the same for both immunity and tolerance.

Analogous results have been obtained with two anti-HLA-DR reagents, DA2 (ref. 23), CA2.06 (ref. 24), and three anti-HLA-DC reagents, SDR 4.1 (ref. 25), HIG 78 (ref. 26) and SG 465 (ref. 27). To confirm that MHC involvement in tolerance induction was not unique to clone HA1.7, a second p20 specific clone HA1.4 was also investigated, and showed similar inhibition by the same anti-class II reagents. Questions of cross-reactivity between monomorphic anti-DR or anti-DC antibodies do not affect the major point of this communication, that recognition of MHC class II antigens is of importance in tolerance induction.

However, the experiments reported here do not resolve whether the class II molecules involved in tolerance induction are present on the same T cell, or adjacent T cells. Our previous results have indicated that cloned T cells can present antigen to other T cells, using an 'anti-idiotypic' (anti-receptor) clone¹⁰.

Studies *in vivo* have attempted to examine whether tolerance to minor transplantation antigens is MHC restricted. Matzinger and Waterfield²⁸ used F₁-parent bone marrow chimaeras, but due to the possibility of antigen reprocessing by F₁ macrophages could not make a definitive assessment. More recently Groves and Singer²⁹ have used parent-F₁ irradiation chimaeras. Their results suggest that tolerance of cytotoxic T-cell precursors involves MHC recognition, in keeping with the *in vitro* clonal analysis reported here.

The observation that T-cell tolerance involves MHC recognition is compatible with either a single or dual T-cell receptor concept. However, it is not consistent with the concept that T-cell immunization involves recognition of antigen in association with MHC, whereas tolerance involves recognition of antigen alone³⁰. If 'positive' induction of a helper clone, to proliferate and help, and 'negative' induction to be tolerant both depend on recognition of the same antigen and MHC product, it is pertinent to ask what factors discriminate between these opposite tendencies. At present antigen concentration and antigen-presenting cells have been shown to be relevant. However, the role of other signals, such as interleukin-1, remain to be elucidated.

We thank J. Knight and S. Smith for technical assistance, Drs P. C. L. Beverley, J. Bodmer, M. Crumpton, S. Goyert, J. Silver, V. van Heyningen for the gift of monoclonal antibodies, Drs N. Green and R. A. Lerner for the gift of synthetic polypeptides of influenza haemagglutinin, Dr R. G. Webster for the gift of haemagglutinin, Dr T. Davies and his colleagues at the National Blood Transfusion Centre, Edgware, for assistance and the Imperial Cancer Research Fund and NIH for financial support.

Received 5 August; accepted 12 December 1983.

1. Zinknagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51–77 (1979).
2. Bodmer, W. F. in *Mammalian Genetics and Cancer* (ed. Russel, E.) 213–240 (Alan R. Liss, New York, 1981).
3. Moller, G. *Immun. Rev.* **66** (1982).
4. Fathman, C. G. & Fitch, F. W. (eds) *Isolation, Characterisation and Utilisation of T Lymphocyte Clones* (Academic, New York, 1982).
5. Feldmann, M. & Schreier, M. H. (eds) *Lymphokines 5* (Academic, New York, 1982).
6. Lamb, J. R., Eckels, D. D., Lake, P., Woody, J. N. & Green, N. *Nature* **300**, 66–69 (1982).
7. Lamb, J. R., Skidmore, B. J., Green, N., Chiller, J. M. & Feldmann, M. *J. exp. Med.* **157**, 1434–1447 (1983).
8. Zanders, E. D., Lamb, J. R., Green, N., Feldmann, M. & Beverley, P. C. L. *Nature* **303**, 456–458 (1983).
9. Lamb, J. R. *et al.* *J. Immun.* **129**, 1465–1470 (1982).
10. Lamb, J. R. & Feldmann, M. *Nature* **300**, 456–458 (1982).
11. Lamb, J. R. *et al.* *Immunology* **50**, 397–405 (1983).
12. Green, N. *et al.* *Cell* **28**, 477–487 (1982).
13. Lamb, J. R. & Green, N. *Immunology* **50**, 659–666 (1983).
14. Shackelford, D. A., Kaufman, J. A., Korman, A. J. & Strominger, J. L. *Immun. Rev.* **66**, 133–187 (1982).
15. Trowsdale, J., Lee, J. & McMichael, A. *Immun. Today* **4**, 31–35 (1983).
16. Lamb, J. R., Woody, J. N., Hartzman, R. J. & Eckels, D. D. *J. Immunol.* **129**, 1465–1470 (1982).
17. Eckels, D. D. *et al.* *Nature* **301**, 716–718 (1983).
18. Schwartz, R. H., Yano, A. & Paul, W. E. *Immun. Rev.* **40**, 153–180 (1978).
19. de Kretzer, Th., Crumpton, M., Bodmer, J. G. & Bodmer, W. F. *Eur. J. Immun.* **12**, 600–606 (1982).
20. de Kretzer, Th., Crumpton, M., Bodmer, J. G. & Bodmer, W. F. *Molecular Biology in Medicine* **1**, 59–76 (1983).
21. Fischer, A., Beverley, P. C. L. & Feldmann, M. *Nature* **294**, 166–168 (1981).
22. Moretta, A., Panaleo, G., Moretta, L., Cerottini, J.-C. & Mingari, M. C. *J. exp. Med.* **157**, 743–754 (1983).
23. Brodsky, F. M., Parham, P., Barnstable, C. J., Crumpton, M. J. & Bodmer, W. F. *Immun. Rev.* **47**, 1–63 (1979).
24. Charron, D. J. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6567–6571 (1979).
25. Rudd, C. E., Bodmer, J. G., Bodmer, W. F. & Crumpton, M. J. *Biochem. Soc. Trans.* (in the press).
26. Guy, K., Van Heyningen, V., Dewar, E. & Steel, C. M. *Eur. J. Immun.* **13**, 156–159 (1983).
27. Goyert, S. M., Shively, J. E. & Silver, J. J. *exp. Med.* **156**, 550–566 (1982).
28. Matzinger, P. & Waterfield, J. D. *Nature* **285**, 492–494 (1980).
29. Groves, E. S. & Singer, A. *J. exp. Med.* **158**, 1483–1497 (1983).
30. Cohn, M. & Epstein, R. *Cell Immun.* **39**, 125–153 (1978).
31. Beverley, P. C. L. & Callard, R. E. *Eur. J. Immun.* **11**, 329–334 (1981).
32. Beverley, P. C. L. *Transplant. clin. Immun.* **11**, 87–94 (1980).
33. Trucco, M. M., Stoker, J. W. & Ceppellini, R. *Nature* **273**, 666–668 (1980).
34. Friedman, S. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **74**, 711–715 (1977).

Chronic benzodiazepine treatment decreases postsynaptic GABA sensitivity

D. W. Gallager, J. M. Lakoski, S. F. Gonsalves & S. L. Rauch

Departments of Psychiatry and Neuroanatomy, Yale University School of Medicine, 34 Park Street, New Haven, Connecticut 06508, USA

Benzodiazepines exert most of their pharmacological effects by a selective facilitation of the postsynaptic actions of GABA^{1–3}. Clinical^{4,5}, behavioural^{6,7} and electrophysiological^{8,9} studies have shown reduced drug response following chronic benzodiazepine administration. We present here electrophysiological evidence for decreased postsynaptic sensitivity to GABA following chronic benzodiazepine administration as measured by the direct iontophoretic application of GABA and serotonin onto serotonergic cells in the midbrain dorsal raphe nucleus (DRN), known to receive GABAergic input^{8–12}. The subsensitivity to GABA was found to be dose dependent and was seen when diazepam administration was three weeks or longer. Further, acute injection of the specific benzodiazepine antagonist¹³, Ro15–1788, was found to reverse rapidly the decrease in

GABA sensitivity observed in chronically diazepam-treated animals without altering GABA sensitivity in vehicle-treated rats. Decreased response to chronic benzodiazepines does not appear to be consistently related to alterations in the number or affinity of receptors for benzodiazepines¹⁴. Our studies of radioligand-binding showed a decrease in the ability of GABA to enhance benzodiazepine binding in cerebral cortical membranes from chronic diazepam-treated animals without significant changes in benzodiazepine binding site density or affinity.

Adult male Sprague-Dawley rats were injected daily with diazepam (DZ: 2.5 or 5 mg kg⁻¹ intraperitoneal) or vehicle (VEH: 40% propylene glycol: 10% ethanol, 1 ml kg⁻¹) for a minimum of 3 weeks prior to electrophysiological recording sessions. Animals were anaesthetized with chloral hydrate 16–24 h after the last injection of DZ or VEH and the single-unit activity of serotonin-containing cells in the DRN was monitored. The DRN was chosen as serotonergic neurones in this nucleus receive GABAergic input^{10–12}; acute benzodiazepines significantly increase the inhibition produced by GABA in the DRN by a postsynaptic mechanism without altering sensitivity to other putative neurotransmitters (serotonin or glycine)¹⁰. Procedures for recording from identified serotonergic neurones in the DRN and the microiontophoretic techniques utilized were as previously described. Pairs of animals were recorded, one from each treatment group in random order until the entire series of recordings were completed (typical traces of spontaneously firing cells from chronic diazepam- or VEH-treated animals are presented in Fig. 1). Throughout the recording series the investigators were blind to treatment history of animals. In these studies, long-term treatment (3 weeks or more) with the 5.0 mg per kg dose resulted in a significant decrease in the sensitivity of serotonergic cells to microiontophoretic GABA (if *I* is ejection current (nA) and *T* is time (s) for DZ: $IT_{50} = 22.2 \pm 2.8$ nA s compared with VEH: $IT_{50} = 10.7 \pm 1.4$ nA s, $P < 0.001$; Fig. 2). The 2.5 mg per kg dose also diminished GABA sensitivity, although the change was not statistically significant. In contrast, iontophoretically applied serotonin was found to be equally effective in both VEH- and diazepam-treated animals suggesting that decreased sensitivity of long-term treated animals is somewhat selective for GABA (Fig. 2). Since variations in basal firing rates have been found to alter drug sensitivities¹⁵, we compared the spontaneous firing rates of DR neurones in DZ- and VEH-treated rats. As previously observed following short-term administration of clinically relevant doses of benzodiazepines^{10,16}, no significant differences in basal firing rates between diazepam and VEH treatments were observed (diazepam: 0.8 ± 0.1 spikes s⁻¹ versus VEH: 1.0 ± 0.1 spikes s⁻¹, $t(67) = 1.71$, not significant). Thus, differences in sensitivity to GABA could not be accounted for by alterations in baseline firing rates of serotonergic cells within the DRN.

To determine whether the subsensitivity to GABA observed following long-term benzodiazepine administration was mediated through a benzodiazepine binding site, the effect of the specific benzodiazepine antagonist¹³, Ro15–1788, was tested in long-term diazepam-treated rats. In these studies, after sampling iontophoretic sensitivity to both serotonin and GABA in a number ($n = 1–4$) of identified serotonin-containing cells within the DRN, sensitivity to a single neurone was repetitively tested both before and after the systemic administration of Ro15–1788 (2 mg per kg, intravenously in suspension with 0.2% carboxymethyl cellulose and 0.1% Tween 80). In long-term diazepam treated animals, Ro15–1788 increased the sensitivity of serotonergic neurones to GABA to the range observed in VEH-treated rats without significantly altering their sensitivity to serotonin (Table 1). In contrast, Ro15–1788 did not alter iontophoretic sensitivities to either GABA or serotonin in VEH-treated animals and, in addition, no significant alteration of baseline firing rate was observed in serotonergic neurones within the DRN following Ro15–1788 administration in either treatment group (Table 1).

Fig. 1 a, Long-term treatment with vehicle. The spontaneous firing of a single serotonergic neurone is plotted as a frequency histogram of spikes per 10 s against time; the duration of drug ejection (serotonin, 5HT; GABA, G) is indicated by the bars above the histogram record with the numbers indicating the iontophoretic current in nanoamperes (nA) used for drug ejection. For each cell tested and for each substance applied iontophoretically an IT_{50} value was determined²⁷ from the digital printout as defined by the ejection current I (nA) applied multiplied by the time (T_{50} , s) required to depress cell firing by 50% from the spontaneous basal firing rate. The response of this serotonergic cell to the microiontophoretic application of GABA and serotonin was measured in a rat which had received daily VEH injections for at least 3 weeks. GABA, ejected with 0.5 nA current for 60 s (bar), inhibited neuronal activity with an IT_{50} value = 10 nA s and serotonin, ejected with 8 nA current for 60 s (bar) also inhibited neuronal activity with an IT_{50} value = 200 nA s. **b**, Long-term diazepam (DZ) treatment. The response of this serotonergic cell to the microiontophoretic application of GABA and serotonin was measured in a rat which had received daily injections of diazepam (5.0 mg per kg) for at least 3 weeks. GABA ejected at currents (0.5 nA) which significantly inhibit spontaneous cell firing in VEH-treated control animals did not significantly depress cell firing in the long-term diazepam-treated animals. When higher ejecting currents were used (3 nA), GABA (60 s; bars) inhibited neuronal activity with an IT_{50} value = 60 nA s. Serotonin, ejected with 4 nA current for 60 s, inhibited cell firing with an IT_{50} value = 155 nA s.

Methods: Serotonin-containing cells in the dorsal raphe nucleus were tentatively identified by a characteristically wide duration action potential (1–2 ms duration, positive-negative wave form), a regular rhythm and a slow firing rate (0.5–3.0 spikes s^{-1}). Spontaneously firing dorsal raphe neurones with these firing pattern characteristics were typically encountered at 4.5–5.5 mm ventral to the level of the exposed tissue on the midline above the DRN. Extracellular recorded spike activity was passed through a high input impedance amplifier and displayed on an oscilloscope; a variable voltage window discriminator was used to assist in single unit discrimination. The electrode signals were then passed into an electronic rate meter triggered by individual spikes. An integrated frequency histogram was graphically recorded in 10-s intervals from the analogue output of the counter; a digital printout of the spike activity was also collected for further statistical analysis. In these studies a five-barrel micropipette was prepared as previously described¹⁰; the protruding centre barrel was filled with a 2 M NaCl 2% Pontamine sky blue dye solution and used for recording (*in vitro* impedance 3–7 M Ω). Side barrels contained solutions of GABA (0.01 M in 0.1 M NaCl, pH 4; 0.001 M in 0.1 M NaCl, pH 4) and serotonin (0.04 M or 0.05 M, pH 4, creatinine sulphate salt) for microiontophoretic application and 4 M NaCl for continuous automatic current balancing; a retaining current of –10 nA was consistently applied to each drug barrel between periods of injection. The duration of drug ejection was controlled to provide a uniform 60-s interval during each microiontophoretic drug application and a constant time interval was maintained between successive, repetitive drug applications. At the conclusion of the recording period the recording site was marked by ejection of a spot of Pontamine sky blue dye. Brain tissue was fixed and examined histologically for localization of the recording site. Localization of the recording site within the DRN was a necessary criterion for inclusion of data in this study.

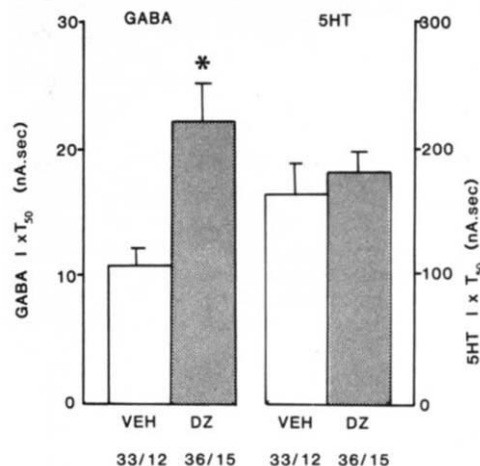
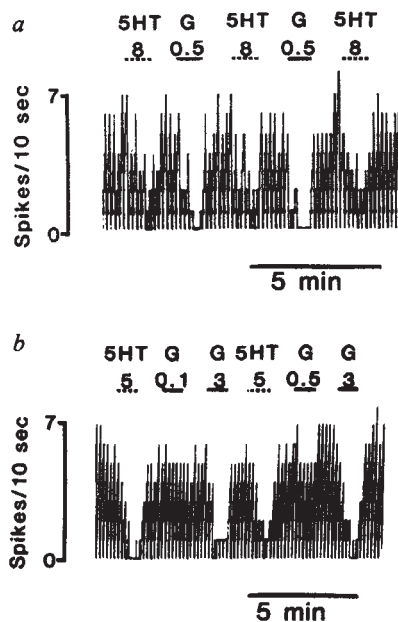


Fig. 2 Sensitivity of spontaneously firing serotonergic cells in the DRN to microiontophoretically applied GABA and serotonin (5HT). Mean IT_{50} values (nA) were determined (see Fig. 1) in paired VEH or 5.0 mg per kg per day diazepam-treated rats from two separate treatment series run 3–4 months apart ($n = 6$ –8 rats per treatment per series). * Overall analysis of variance showed a significant effect of DZ treatment on GABA IT_{50} , $F(1,70) = 13.32$, $P < 0.001$, with no significant effect on 5HT IT_{50} , $F(1,64) = 0.57$, not significant. Numbers below each bar show no. of cells/no. of rats.

These data indicate that serotonergic neurones in the DRN develop a functional subsensitivity to GABA following chronic exposure to diazepam. Since drug treatment was discontinued 14–16 h before recording sessions, the subsensitivity to GABA observed in chronic diazepam-treated rats could be related either to development of tolerance or to withdrawal of the drug. The benzodiazepine antagonist, Ro15-1788, has been reported to induce benzodiazepine withdrawal responses in a number of experimental studies^{17–19}. In our studies, Ro15-1788 administration rapidly reversed GABA subsensitivity in long-term DZ-treated rats. These data therefore indicate that the decreased sensitivity to GABA is not a withdrawal response, a response which would be predicted to be enhanced rather than reversed by Ro15-1788. In addition, because long-term diazepam treatment results in significant physiological changes in the CNS, these data corroborate evidence that tolerance is not due solely to pharmacokinetic differences in chronic drug disposition⁵. Since subsensitivity to the microiontophoretic application of GABA following long-term diazepam treatment is not seen following a single 24-h exposure to diazepam (diazepam: $IT_{50} = 11.8 \pm 0.4$ nA s versus VEH: $IT_{50} = 11.2 \pm 0.3$ nA s, $t(10) = 0.29$, not significant) and, importantly, is in fact opposite the facilitation of GABA responses immediately following diazepam administration¹⁰, the data suggest that such changes cannot be due simply to the acute presence of diazepam in chronically treated animals.

To further address this issue, levels of diazepam and active metabolites remaining in brain tissue at the time of investigation were determined using a radioreceptor assay²⁰. Eight brainstems were selected at random from long-term diazepam-treated rats used in these studies. The assay used, which was able to detect a minimum level of 3 pmol per assay of diazepam, failed to detect the presence of active drug in 6 out of 8 tissue samples. In tissue from the two remaining animals, less than 0.01% of the 5.0 mg per kg daily dose of diazepam or its active metabolites was measured. In view of these findings, there is a low probability that any significant levels of diazepam were present in brain tissue at the time of the electrophysiological or biochemical analysis. Thus, although repeated occupation of benzodiazepine sites is required for these effects, the continuing presence of pharmacologically active concentrations of drug in brain²¹ does not seem necessary for the development of subsensitivity.

Table 1 Antagonism of subsensitivity to GABA by Ro15-1788

Treatment	Microiontophoretic sensitivity [mean IT_{50} (nA s) $\pm$ s.e.m.]*						Basal firing rate [spikes per 10 s]		
	GABA			Serotonin			Pre-Ro	(n)	Post-Ro
Chronic VEH	15.7 $\pm$ 5.7	(7)	10.8 $\pm$ 4.5	238 $\pm$ 70	(5)	222 $\pm$ 62	14.1 $\pm$ 1.6	(7)	13.6 $\pm$ 1.6
Chronic diazepam	32.0 $\pm$ 6.3	(8)	16.2 $\pm$ 2.2†	199 $\pm$ 47	(8)	173 $\pm$ 29	8.9 $\pm$ 1.1	(8)	8.4 $\pm$ 1.2

* Microiontophoretic sensitivity determined from average of two or three successive 60-s applications of serotonin and GABA (Fig. 1) immediately preceding and following intravenous injection of 2 mg per kg Ro15-1788.

† $P < 0.05$, paired Student's *t*-test (Pre versus Post-Ro15-1788).

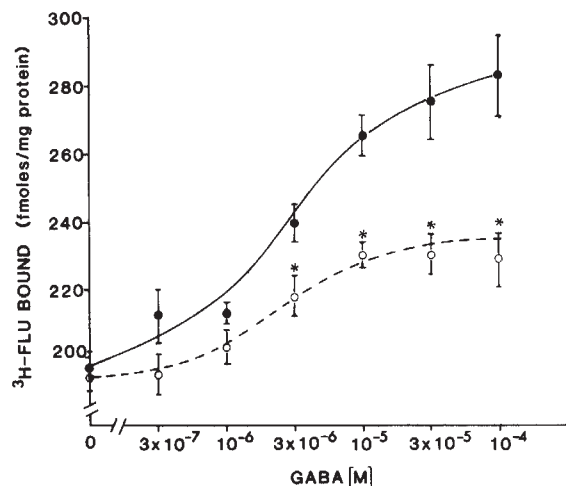


Fig. 3 GABA stimulation of ^3H -flunitrazepam binding in cerebral cortical membranes prepared from rats treated daily for 21 days with vehicle (○) or diazepam (5.0 mg per kg) (●). Briefly, crude brain membranes were extensively washed (5×) to remove endogenous GABA²⁷. Immediately prior to the binding assay procedure the tissue was resuspended in 10 vol of 50 mM Tris HCL, pH 7.0 buffer and dispersed with a Polytron (setting 8.5 for 15 s; Brinkman). Protein concentration was determined for each sample²⁹ and the data expressed per mg protein. Binding to cortical membranes was measured in an incubation volume of 500 μl (100 μl tissue homogenate, 350 μl Tris HCL buffer or drug dissolved in the Tris buffer, and 50 μl of ^3H -flunitrazepam (specific activity = 87.6 nCi mmol⁻¹, New England Nuclear) at a concentration of 0.5 nM as previously described²⁷. Incubation time was for 30 min at 37 °C and terminated by rapid filtration through Whatman GF/B filters followed by 3×6 ml rinses with ice-cold Tris buffer. Various concentrations of GABA were added to the incubation medium immediately prior to the addition of ^3H -flunitrazepam (1 nM). Each point ($\pm$ s.e.m.) is the mean of specifically bound ^3H -flunitrazepam determined in three separate experiments performed in quadruplicate. * Significantly different ($P < 0.05$) from vehicle rat values according to paired Student's *t*-test.

To investigate whether receptor alterations accompany the observed changes in GABA sensitivity, we examined benzodiazepine binding and the ability of GABA to enhance benzodiazepine binding in long-term treated animals. There are conflicting reports of alterations in benzodiazepine binding following a wide variety of chronic dose regimens¹⁴. Since membranes prepared from punches including the DRN region yielded insufficient quantities of tissue for these studies, extensively washed cerebral cortical membranes (Fig. 3) were used. However, previous studies have shown that GABAergic modulation of benzodiazepine binding is similar to effects on membranes prepared from the DRN region²¹. Our treatment regimen, which results in electrophysiologically measurable subsensitivity to GABA in the DRN, did not significantly alter either benzodiazepine receptor number or affinity as measured by saturation binding of ^3H -flunitrazepam (^3H -FLU, 0.5–16 nM, 37 °C) to fresh extensively washed cerebral cortical membranes (diazepam: B_{max} 3.9 $\pm$ 0.3 pmol per mg protein, K_D = 21.8 $\pm$

7 nM versus, VEH: B_{max} 3.7 $\pm$ 0.3 pmol per mg protein, K_D = 21.5 $\pm$ 0.6 nM). In cortical membranes from both VEH- and diazepam-treated animals, GABA was found to enhance ^3H -FLU binding with similar EC_{50} values (VEH: 4 μM GABA versus diazepam: 3.5 μM GABA). However, the magnitude of the GABA enhancement was found to be significantly reduced in membranes from chronic diazepam-treated rats as compared to tissue from VEH-treated rats (VEH: 43.0 $\pm$ 5.0% versus diazepam 21.5 $\pm$ 2.3%; Fig. 3). Decreases in the ability of GABA to enhance diazepam binding have also been observed both in adult mice²² and in developing rats following prenatal exposure to benzodiazepines²³. These data, obtained using ligand-binding techniques, are, therefore, consistent with electrophysiological evidence for subsensitivity to GABA following chronic benzodiazepine administration.

Our receptor binding data suggest that chronic exposure of the receptor to DZ alters the sensitivity to GABA and that such changes are not accompanied by measurable alterations in benzodiazepine receptor number or affinity. Our physiological data indicate that a change in the functional responsivity to GABA in the GABA–benzodiazepine macromolecular complex (or a more complex change in the interactions between various elements of the GABA postsynaptic complex) has occurred following long-term DZ administration. Since benzodiazepines appear to act as allosteric modifiers in the GABA receptor complex^{24–27}, our data would indicate that chronic exposure to a modifier could result in decreased function (subsensitivity) of the entire GABA–benzodiazepine receptor complex. Although the present studies do not address the molecular basis for such changes in functional responsivity to GABA, decreases in GABA sensitivity could be one biological mechanism responsible for the development of tolerance to the benzodiazepines.

This work was supported by USPHS grant MH-14276, a grant from the Esther A. and Joseph Klingenstein Fund to D.W.G. and the State of Connecticut. We thank Ruth Vergason for secretarial assistance.

Received 8 August; accepted 7 December 1983.

- Costa, E. & Guidotti, A. A. *Rev. Pharmac. Tox.* **19**, 531–545 (1979).
- Haefely, W. *et al.* In *GABA-Neurotransmitters* (eds Krosgaard-Larsen, P., Scheel-Krueger, J. & Kofod H.) 357–375 (Munksgaard, Copenhagen, 1979).
- Tallman, J. F., Paul, S. M., Skolnick, P. & Gallager, D. W. *Science* **207**, 274–281 (1980).
- Browne, T. R. In *Pharmacology of Benzodiazepines* (eds Usdin, E., Skolnick, P., Tallman, J., Greenblatt, D. & Paul, S.) 329–337 (Macmillan, London, 1983).
- Greenblatt, D. J. & Shader, R. J. *Drug Metabolism. Rev.* **8**, 13–28 (1978).
- Rosenberg, H. C. & Chiu, T. H. *Eur. J. Pharmac.* **70**, 453–460 (1981).
- Hironaka, T., Fuchino, K. & Fujii, T. *Japan J. Pharmac.* **33**, 95–102 (1983).
- Waterhouse, B., Moises, H., Yeh, H., Geller, H. & Woodward, D. *J. Pharmac. exp. Ther.* (in the press).
- Sher, P., Study, R., Mazzetta, J., Barker, J. & Nelson, P. *Brain Res.* **268**, 171–176 (1983).
- Gallager, D. *Eur. J. Pharmac.* **49**, 133–143 (1978).
- Belin, M. *et al. Brain Res.* **170**, 279–297 (1979).
- Nanopoulos, D., Belin, M., Maitre, M., Vincendon, G. & Pujol, J. *Brain Res.* **232**, 375–389 (1982).
- Hunkeler, W. *et al. Nature* **209**, 514–516 (1981).
- Braestrup, C. & Nielsen, M. In *Handbook of Psychopharmacology* (eds Iversen L. L., Iversen S. D. & Snyder S. H.) 285–384 (Plenum, New York, 1983).
- Andrade, R., VanderMaelen, C. & Aghajanian, G. *Eur. J. Pharmac.* **291**, 161–169 (1983).
- Trulsson, M., Preussler, D., Howell, G. & Frederickson, C. *Neuropharmacology* **21**, 1045–1050 (1982).
- Cumin, R., Bonetti, E., Scherschlicht, R. & Haefely, W. *Experientia* **38**, 833–834 (1982).
- Lukas, S. E. & Griffiths, R. R. *Science* **217**, 1161–1163 (1982).
- McNicholas, L. & Martin, W. *Fedn Proc.* **41**, 1639 (1982).
- Lund, J. *Scand. J. clin. Lab. Invest.* **41**, 275–280 (1981).
- Gallager, D. & Tallman, J. *Neuropharmacology* (in the press).
- Braestrup, C. & Nielsen, M. *Adv. Biosci.* **31**, 221–227 (1981).

23. Massotti, M., Alleva, F. R., Balazs, T. & Guidotti, A. *Neuropharmacology* **19**, 951-956 (1980).
 24. Braestrup, C., Schmieden, R., Neef, G., Nielsen, M. & Petersen, E. N. *Science* **216**, 1241-1243 (1982).
 25. Costa, E., Corda, M., Epstein, B., Forchetti, C. & Guidotti, A. In *Benzodiazepines: Molecular Biology to Clinical Practice* (ed. Costa, E.) 117-146 (Raven, New York, 1983).

26. Olsen, R. W. A. *Rev. Pharmac. Tox.* **22**, 45-77 (1982).
 27. Tallman, J. F., Thomas, J. W. & Gallager, D. W. *Nature* **274**, 383-385 (1978).
 28. Gallager, D. W. & Bunney, W. E. *Naunyn-Schmiedeberg's Arch. Pharmac.* **307**, 129-133 (1979).
 29. Lowry, O. H., Rosebrough, M. J., Farr, A. L. & Randall, R. L. *J. biol. Chem.* **193**, 265-275 (1951).

Voltage-dependent calcium channels from brain incorporated into planar lipid bilayers

Mark T. Nelson*, Robert J. French†
 & Bruce K. Krueger*

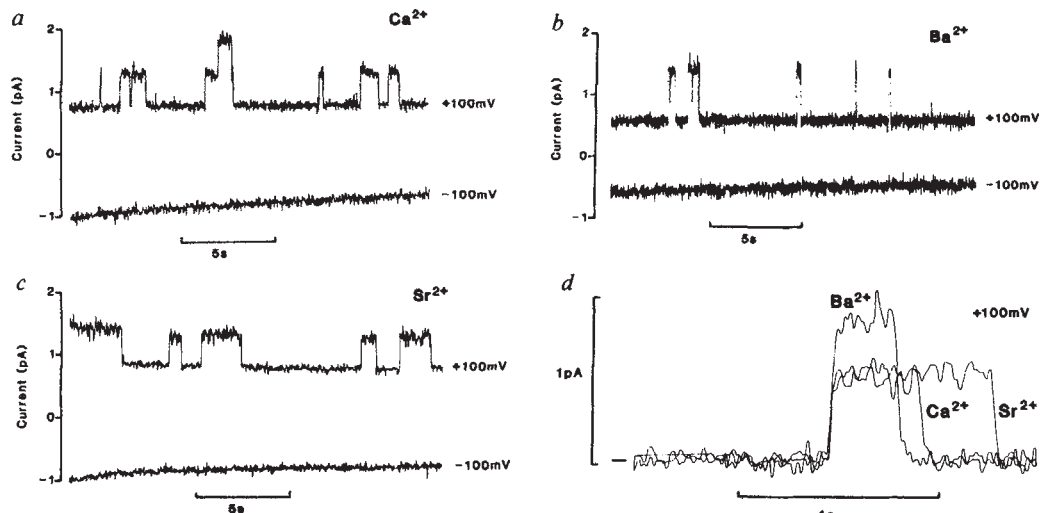
Departments of Physiology* and Biophysics†, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA

Many important physiological processes, including neurotransmitter release and muscle contraction¹⁻³, are regulated by the concentration of Ca^{2+} ions in the cell. Levels of cytoplasmic Ca^{2+} can be elevated by the entry of Ca^{2+} ions through voltage-dependent channels which are selective for Ca^{2+} , Ba^{2+} and Sr^{2+} ions⁴⁻¹⁴. We have measured currents through single, voltage-dependent calcium channels from rat brain that have been incorporated into planar lipid bilayers. Channel gating was voltage-dependent: membrane depolarization increased the channel open times and decreased the closed times. The channels were selective for divalent cations over monovalent ions. The well-known calcium channel blockers, lanthanum and cadmium, produced a concentration-dependent reduction of the apparent single-channel conductance. Contrary to expectations¹⁴, the nature of the divalent cation carrying current through the channel affected not only the single-channel conductance, but also the channel open times, with mean open times being shortest for barium.

The median eminence was chosen as a potential source of Ca^{2+} channels because it is rich in nerve terminals which secrete releasing and inhibiting hormones that travel to the pituitary by means of the hypophyseal portal system¹⁵. Channels were incorporated into bilayers from membrane vesicles that were made from rat brain median eminences using a method similar to that used for synaptosomes¹⁶ or membrane vesicles¹⁷ from whole brain. Planar bilayers containing phosphatidylethanolamine (33 mg ml⁻¹, bovine brain; Avanti Polar Lipids) and phosphatidylserine (26 mg ml⁻¹, bovine brain; Avanti) in decane were painted across a 100-250 µm hole in a polystyrene or Lexan partition separating two chambers^{16,17}, both (except where noted) containing (in mM) one of 250 CaCl_2 or SrCl_2 or BaCl_2 with 7.5 HEPES, 0.11 CaCl_2 , 0.075 MgCl_2 , 0.038 EGTA, pH 7.0 (that is, the divalent cation equilibrium potential was normally 0 mV). The *cis* side of the bilayer was defined as that side of the bilayer exposed to the membrane vesicles; the opposite (*trans*) side was held at virtual ground. Because, as inferred by their voltage dependence, the channels incorporated in the bilayer oriented with the extracellular ends of the channels facing the *trans* side and their cytoplasmic ends facing the *cis* side, voltages reported here correspond to the usual cellular convention (inside minus outside).

Within minutes following the addition of the membrane preparation to the *cis* side in the presence of 250 mM CaCl_2 , BaCl_2 or SrCl_2 , stepwise current fluctuations of 0.5 pA (Ca^{2+} , Sr^{2+}) or 0.85 pA (Ba^{2+}) were observed at +100 mV (Fig. 1a-c). The bottom current records in Fig. 1a-c show that fluctuations in

Fig. 1 a, Single-channel current fluctuations with Ca^{2+} . At +100 mV (top record), the current fluctuations were 0.5 pA in symmetric divalent cation solutions. In this experiment there were at least two channels in the bilayer. The holding potential was then changed to -100 mV (bottom record). At -100 mV, current fluctuations were not observed. The difference in current between -100 mV and +100 mV with all channels closed represents the bare bilayer conductance, about 7.5 pS. b, Single-channel current fluctuations with Ba^{2+} . At +100 mV, current fluctuations of about 0.85 pA with symmetric divalent cations were measured. At -100 mV, current fluctuations were not seen. c, Single-channel current fluctuations with Sr^{2+} . At +100 mV, current fluctuations of about 0.5 pA with symmetric divalent cations were measured. At -100 mV, current fluctuations were not observed. Hole diameter: 100 µm. d, Comparison of single-channel currents with Ba^{2+} , Ca^{2+} and Sr^{2+} at +100 mV. Current records taken in the presence of each ion were superimposed. In all cases, symmetric 250 mM divalent cations were present. **Methods:** Single-channel current fluctuations. The membranes were prepared from 10-15 rat brain median eminences, as described in ref. 16 with the omission of the sucrose gradient step. The membranes were suspended in 0.4 M sucrose and could be stored at -75 °C for at least 2 months without affecting the results. Identical results were obtained with both fresh and frozen preparations. The thawed synaptosomes in 0.4 M sucrose were then sonicated (probe type, Braun model 1510) for 20 s immediately before use. Addition of this membrane vesicle suspension (final concentration 0.1-0.5 µg protein ml⁻¹) resulted in stepwise current fluctuations^{16,17}. All experiments were conducted at room temperature (19-25 °C). The side to which membrane vesicles were added was designated the *cis* side; the opposite (*trans*) side was held at virtual ground. Positive currents shown in the figures reflect positive charges moving outward, that is, from the *cis* side to the *trans* side. In all figures shown in this report except Fig. 2c, the divalent cation equilibrium potential was 0 mV. The cut-off frequency of the current to voltage converter was about 600 Hz. All experiments were recorded on FM tape at 600 Hz and, unless otherwise noted, records were filtered at 100 Hz on playback. Unless stated otherwise, the diameter of the holes onto which the bilayers were painted was 250 µm. The slow change in current seen in Figs 1a and c at -100 mV reflected the residual relaxation of the capacitive current following the voltage step from +100 mV to -100 mV which preceded the record shown.



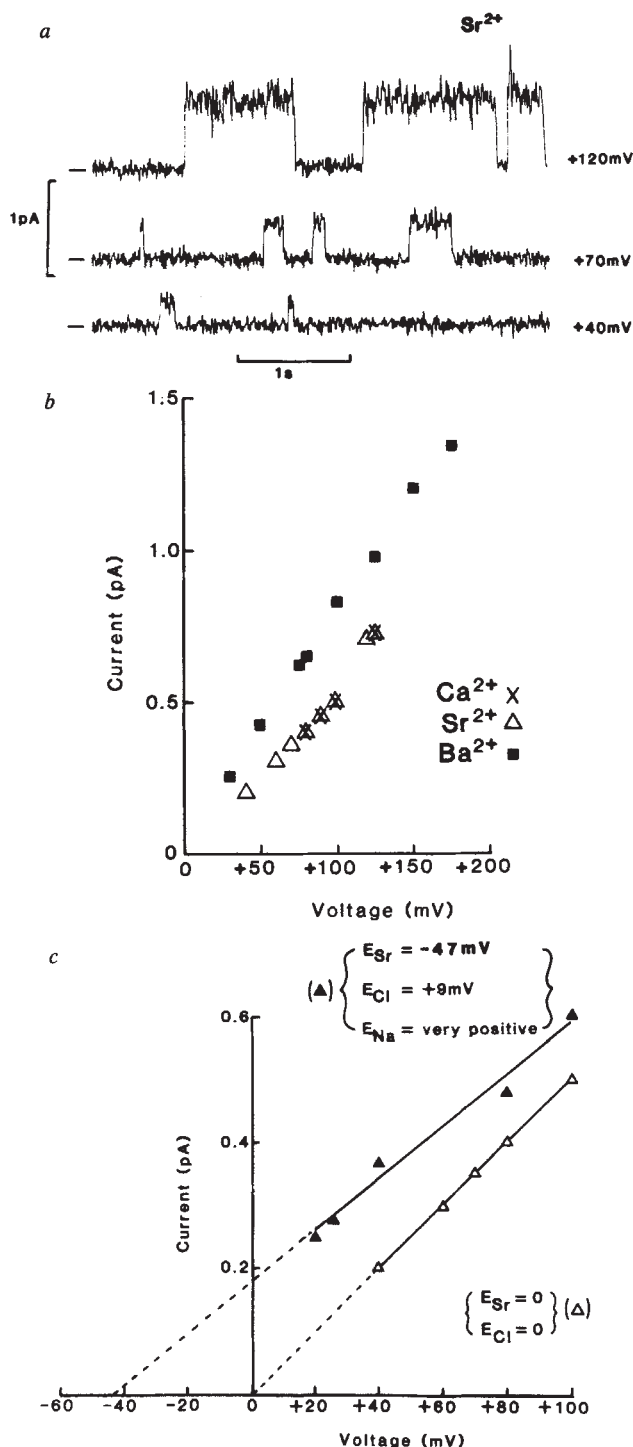


Fig. 2 Current-voltage relations for single calcium channels. *a*, Effects of voltage on the magnitude of single-channel currents. Records were taken on the same membrane with symmetrical 250 mM SrCl_2 . The solid horizontal bars in each record indicate the current level with all channels closed. Hole diameter: 100 μm . Above +100 mV, there is a pronounced increase in current fluctuations in the open state of the channel. *b*, Current-voltage relationship for single calcium channels with Ca^{2+} , Ba^{2+} and Sr^{2+} . Symmetrical 250 mM divalent cations, that is, the zero-current potential and divalent cation equilibrium potential were at 0 mV. Single-channel conductances for Ca^{2+} , Ba^{2+} and Sr^{2+} were 5 pS, 8.5 pS and 5 pS, respectively. *c*, Single-channel current-voltage relations with symmetrical 250 mM SrCl_2 (Δ , same data as in *b*) and with 357 mM NaCl, 6.1 mM SrCl_2 on the *trans* side and 250 mM SrCl_2 on the *cis* side ($\blacktriangle$). The equilibrium potentials (calculated from concentrations) for these experiments were: Δ (symmetric), $E_{\text{Sr}} = E_{\text{Cl}} = 0 \text{ mV}$; $\blacktriangle$ (asymmetric), $E_{\text{Sr}} = -47 \text{ mV}$, $E_{\text{Cl}} = +9 \text{ mV}$, $E_{\text{Na}} = \text{nominal plus infinity}$. The slopes of the linear regression (solid lines) were 5 pS (open triangles) and 4.2 pS (filled triangles).

current were not seen for any of the divalent cations at -100 mV . The current leak through the bilayer in the absence of any channel passing current was symmetrical at $\pm 100 \text{ mV}$ and corresponded to a 'naked bilayer' conductance of about 7.5 pS. Figure 1*d* shows superimposed single channel current steps for Ca^{2+} , Ba^{2+} and Sr^{2+} at +100 mV. Strontium and barium have been found to carry current through all Ca^{2+} channels studied, so that the ability of Sr^{2+} or Ba^{2+} to substitute for Ca^{2+} in maintaining Ca^{2+} currents is an important criterion used in identifying Ca^{2+} channels³. Figure 1 shows first that with equimolar concentrations of the three divalent cations the efficacy in carrying current at +100 mV was $\text{Ba} > \text{Ca} = \text{Sr}$. This agrees with reports for Ca^{2+} channels in other cell types¹⁻¹¹. Second, clear, well-defined current steps were seen for all three divalent cations at +100 mV, while current fluctuations were not seen at -100 mV . Thus, the gating of these single channels is voltage-dependent. Since single-channel current fluctuations continued to be observed for many seconds following a voltage step, the channels under these experimental conditions do not completely inactivate (compare refs 9-11). Thus far, with the conditions described in this paper, we have been able to incorporate Ca^{2+} channels on virtually every attempt (>50) and channels can frequently be studied for periods of 30 to 60 min following incorporation.

Single-channel current-voltage relations for these cations are shown in Fig. 2. Figure 2*a* shows representative single channel records at three potentials with Sr^{2+} as the current carrier. Figure 2*b* shows that the single-channel current with either Ca^{2+} , Ba^{2+} or Sr^{2+} as the current carrier was proportional to applied voltage over a range of about 145 mV. The single-channel conductances with Ca^{2+} , Ba^{2+} and Sr^{2+} as the current carriers were 5 pS, 8.5 pS and 5 pS, respectively, indicating that with symmetrical 250 mM divalent cations, Ba^{2+} moves through the channel about 1.7 times faster than Ca^{2+} or Sr^{2+} . At potentials above +100 mV, open channel current noise with all three divalent cations increased substantially (see Fig. 2*a*). Reducing the Ca^{2+} concentration from 250 to 150 mM had no effect on the single-channel conductance, suggesting that at the levels of cations used in these experiments, the Ca^{2+} channels were saturated with respect to the current carrier^{3,9,11}. The use of symmetric solutions eliminates surface potential differences, and channel saturation eliminates variation in the degree of occupancy by the permeant ions. Thus, the different single-channel conductances should reflect differences among the limiting rates of transport of these ions through the channels. Negative to +30 mV, the single-channel currents were too small and too brief to measure. Patch clamp studies on intact cultured rat heart muscle and clonal rat pituitary cells suggest that single-channel Ba^{2+} currents increase linearly over a 50-60 mV range^{4,5} and that Ba^{2+} carries more current than Ca^{2+} through single channels in *Helix* neurones⁶.

Direct evidence of selectivity for divalent cations over monovalent cations is shown in Fig. 2*c*. In this experiment, the *trans* SrCl_2 solution was replaced with an isotonic solution of NaCl. Although the channels were open long enough to be measured only at potentials more depolarized than +20 mV, the single channel current-voltage relationship was linear and could be extrapolated to a potential (-44 mV) that was very close to the divalent cation equilibrium potential (E_{Sr} , -47 mV). By contrast, the chloride equilibrium potential (E_{Cl}) was +9 mV and the sodium equilibrium potential (E_{Na}) was nominally plus infinity. This result rules out the possibility that these channels select for Na^+ or Cl^- . In other experiments, a high concentration (200 mM) of KCl was added to the *cis* side with 150-250 mM CaCl_2 or BaCl_2 present on both sides. In these experiments (results not shown), there was no change in the single channel current at any potential. These results also support the conclusion that these channels are selective for divalent cations over monovalent ions.

As was illustrated in Fig. 1*a-c*, the channels are more likely to be open at positive potentials. We have also analysed the effect of changing potential on the single-channel open and

closed times, which reflect the probabilities of closing and opening. Using either Ca^{2+} , Ba^{2+} or Sr^{2+} as the current carrier, decreasing the membrane potential from +100 mV to +50 mV caused the channel open times to decrease, and the closed times (intervals between openings) to increase (data shown for Ba^{2+} , Fig. 3). At +100 mV, the mean open time of the channel with Ba^{2+} as the current carrier was about 127 ms (Fig. 3*b*). Reducing the potential to +75 mV and then to +50 mV decreased the mean open time of the channel to 76 ms and to 29 ms. If the mean open time continues to decrease as the membrane potential is reduced further, then the channel open times should have been in the millisecond range around 0 mV, a value similar to those obtained in patch clamp experiments on other preparations⁴⁻⁶. Reducing the potential not only decreased channel open times but it also increased the channel closed times (Fig.

3*c*). From +100 mV to +50 mV, the closing rate increased 4.4-fold and the opening rate decreased 2.5-fold indicating that the probability of a channel being open changed approximately 10-fold for this 50 mV change in membrane potential, a value similar to that found for Ba^{2+} -conducting channels in cultured heart muscle cells⁴.

Channel open times were affected by the species of ion carrying the current through the channel as well as by voltage (contrast ref. 22). Throughout the range of potentials studied, channels with Ba^{2+} carrying the current had a shorter mean open time than with Ca^{2+} or Sr^{2+} as current carriers (Fig. 3*d*). At +100 mV, the single-channel mean open times with Ba^{2+} , Sr^{2+} and Ca^{2+} were 127 ms, 385 ms and 454 ms, respectively. Thus, there is a parallel between the sequence of mean open times, and the sequence of mean transit times for the permeant ions (0.4 μs for Ba^{2+} ; 0.7 μs for Ca^{2+} and Sr^{2+} , calculated from the unit current at +100 mV). A similar relationship can be derived from open times and unit conductances observed when different alkali cations carry current through the acetylcholine receptor channel^{18,19}. In addition, the order of the mean open times ($\text{Ba}^{2+} < \text{Sr}^{2+} < \text{Ca}^{2+}$) is the same as the order of apparent affinities of these ions for Ca^{2+} channels in rat brain synaptosomes⁹. These results are consistent with the general idea that occupancy of a channel by a permeant ion can impede channel closing, a conclusion also reached in studies of squid axon potassium channels²⁰.

Calcium entry into a wide variety of cells is mediated by voltage-activated channels that are permeable to the cations Ca^{2+} , Ba^{2+} and Sr^{2+} (see Fig. 1; for review see ref. 3). The exact mechanism of ion permeation is unknown, but one important step may be the association of Ca^{2+} with a binding site in

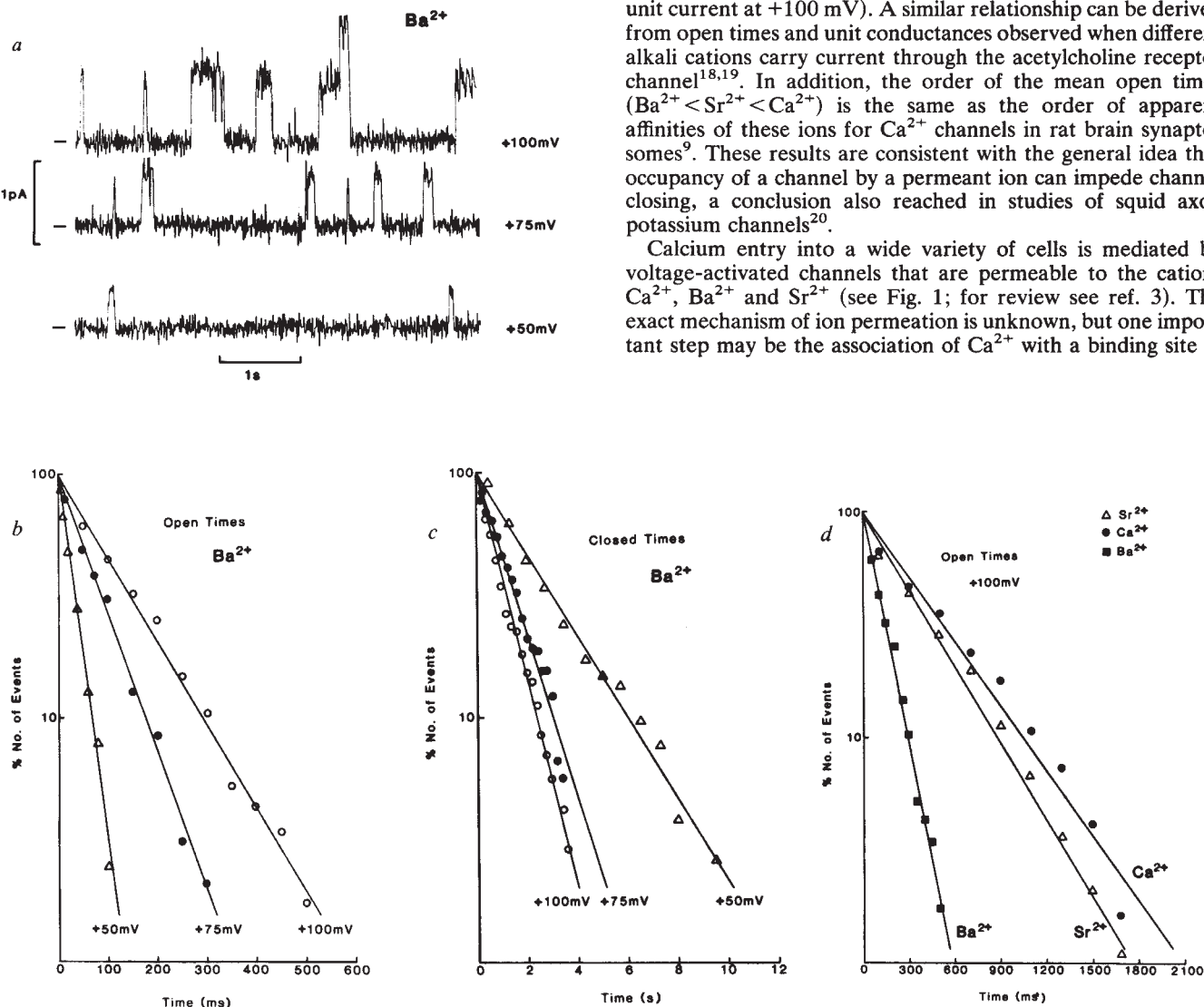


Fig. 3 Voltage-dependence of single-channel open times and closed times. *a*, Original record of single-channel current fluctuations with Ba^{2+} as the current carrier at +50 mV, +75 mV and +100 mV. The solid bars show the level at which all the channels were closed. *b*, Effect of voltage on the distribution of single-channel open times. Open-time distributions were determined from the same experiment shown in 3*a*. The open times were determined from single-level fluctuations that terminated in a closure; multiple-level fluctuations were rare and were not counted. The distributions of open times at +50 mV, +75 mV, +100 mV could be described in each case by a single exponential. Plotted on a logarithmic scale on the ordinate is the percentage of events that lasted at least as long as the time indicated on the abscissa. The total number of events, the slopes and correlation coefficients of the linear regression lines through the log-transformed data were at +50 mV, 142, -34.7 s^{-1} , -0.993 ; at +75 mV, 77, -13.1 s^{-1} , -0.998 ; and at +100 mV, 75, -7.9 s^{-1} , -0.999 . *c*, Effect of voltage on the distribution of times between openings from the same experiment. The distributions of closed times at +50 mV, +75 mV and +100 mV could also be described by single exponentials. The total number of events, the slopes and correlation coefficients of the linear regression lines through the log-transformed data were at +50 mV, 142, -0.38 s^{-1} , -0.993 ; at +75 mV, 77, -0.76 s^{-1} , -0.985 ; and at +100 mV, 71, -0.94 s^{-1} , -0.997 . *d*, Effects of Ba^{2+} , Ca^{2+} and Sr^{2+} on the distribution of single-channel open times at +100 mV. The total number of events and the slopes and correlation coefficients of the linear regression lines through the log-transformed data were with Ba^{2+} , 75, -7.9 s^{-1} , -0.999 ; with Sr^{2+} , 124, -2.6 s^{-1} , -0.996 ; and with Ca^{2+} , 82, -2.2 s^{-1} , -0.985 .

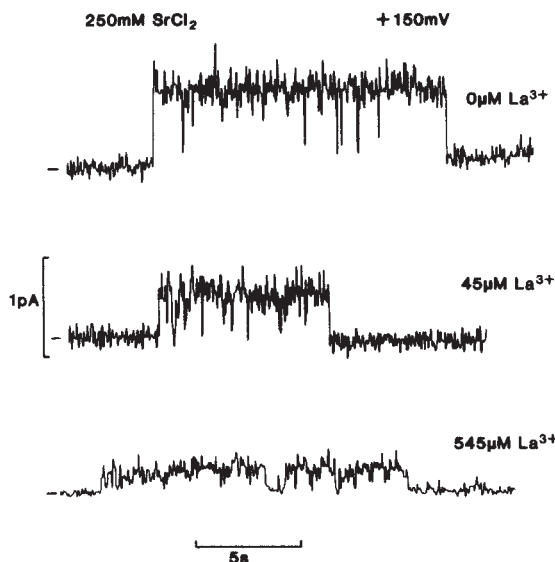


Fig. 4 Block of single-channel strontium currents by lanthanum. Representative single channel current steps in the presence of 250 mM SrCl_2 in 0, 45 and 545 μM LaCl_3 on the *cis* side. The top and middle records were filtered at 100 Hz and the bottom record was filtered at 20 Hz on playback.

the channel. There is good evidence that in many types of cells, lanthanum and a variety of transition metals block Ca^{2+} movement through the channels by competing with Ca^{2+} for this site. To date, there are no data at the single channel level on the mechanism of block by inorganic ions. We found that lanthanum reduces the single channel current in a dose-dependent manner (Fig. 4). In the presence of 250 mM Sr^{2+} , 145 μM La^{3+} reduced the single channel current by 50%. In addition to lanthanum, cadmium and manganese also reduced the single channel currents, with the order of potency being: $\text{La}^{3+} > \text{Cd}^{2+} \gg \text{Mn}^{2+}$.

The selectivity among Ca^{2+} , Ba^{2+} and Sr^{2+} , and for divalent over monovalent cations, the small single-channel conductances, the voltage dependence, and the block by lanthanum and cadmium, reported here, are considered to be properties common to all Ca^{2+} channels¹⁻³. Based on these criteria, these channels closely resemble voltage-dependent calcium channels that have been studied in a variety of excitable tissues. Of several unresolved questions about Ca^{2+} channels, the following appear to be especially well-suited for investigation using the bilayer system: (1) Can permeant divalent cations move through the channels in both directions? In the presence of physiological Ca^{2+} gradients, current through Ca^{2+} channels is inward. Our results probably reflect an outward movement of divalent cations through the open channels. Bidirectional movement could be verified in the presence of the appropriate divalent cation concentration gradients to allow measurable currents at potentials where the channels are open. (2) Are Ca^{2+} channels modulated directly by neurohormones (for example, catecholamines) or via a cyclic AMP-protein phosphorylation cascade^{1,2}? This hypothesis could be tested in the planar bilayer system by adding either the neurohormones or the constituents necessary for protein phosphorylation. (3) How do clinically important 'Ca-antagonists' affect single Ca^{2+} channels²¹? The side of the channels at which these agents act and the reversibility of their action could easily be investigated. Finally, the planar bilayer system may also provide the opportunity to compare Ca^{2+} channels from different tissues, such as heart and smooth muscle, under well-defined conditions.

We thank W. J. Lederer and D. A. Nachshen for helpful discussions, M. P. Blaustein, M. Cannell, E. Marban and J. F. Worley III for comments on the manuscript, E. Bamberg and P. Luger for insights into the planar bilayer system and ion

channel function, K. Gregerson and M. Selmanoff for introducing us to the median eminence preparation and D. Darragh for preparing the figures. This work was supported by a Grant-in-aid from the American Heart Association, Maryland Affiliate (M.T.N.), and the Alexander von Humboldt-Stiftung (M.T.N.), the Bressler Research Fund of the University of Maryland (R.J.F. and B.K.K.) and contract DAMD17-82-C-2188 from the US Army Medical Research and Development Command (B.K.K. & R.J.F.).

Received 21 July; accepted 29 November 1983.

1. Tsien, R. W. *A. Rev. Physiol.* **45**, 341-358 (1983).
2. Reuter, H. *Nature* **301**, 569-574 (1983).
3. Hagiwara, S. & Byerly, L. A. *Rev. Neurosci.* **4**, 69-125 (1981).
4. Reuter, H., Stevens, C. F., Tsien, R. W. & Yellen, G. *Nature* **297**, 501-504 (1982).
5. Hagiwara, S. & Ohmori, H. *J. Physiol., Lond.* **336**, 649-661 (1983).
6. Lux, H. D. & Nagy, K. *Pflügers Arch. ges. Physiol.* **391**, 252-254 (1981).
7. Fenwick, E. M., Marty, A. & Neher, E. *J. Physiol., Lond.* **331**, 599-635 (1982).
8. Brown, A. M., Camerer, H., Kunze, D. L. & Lux, H. D. *Nature* **299**, 156-158 (1982).
9. Nachshen, D. A. & Blaustein, M. P. *J. gen. Physiol.* **79**, 1065-1087 (1982).
10. Llinas, R., Steinberg, I. Z. & Walton, K. *Biophys. J.* **33**, 289-322 (1981).
11. Hagiwara, S., Fukuda, J. & Eaton, D. C. *J. gen. Physiol.* **63**, 564-578 (1974).
12. Ehrlich, B. E., Finkelstein, A., Forte, M. & Kung, C. *Biophys. J.* **41**, 293a (1983).
13. Venter, J. C. *et al. J. biol. Chem.* **258**, 9344-9348 (1983).
14. Hagiwara, S. & Ohmori, H. *J. Physiol., Lond.* **331**, 231-252 (1982).
15. Kobayashi, H. & Matzu, T. in *Frontiers in Neuroendocrinology* (eds Ganong, W. F. & Martini, L.) 3-46 (Oxford University Press, 1969).
16. Nelson, M. T., Roudna, M., & Bamberg, E. *Am. J. Physiol.* **245**, C151-156 (1983).
17. Krueger, B. K., Worley III, J. F. & French, R. J. *Nature* **303**, 172-175 (1983).
18. Marchais, D. & Marty, A. *J. Physiol., Lond.* **297**, 9-45 (1979).
19. Gage, P. W. & Van Helden, D. J. *Physiol., Lond.* **288**, 509-528 (1979).
20. Swenson Jr., R. P. & Armstrong, C. M. *Nature* **291**, 427-429 (1981).
21. Lee, K. S. & Tsien, R. W. *Nature* **302**, 790-794 (1983).
22. Saimi, Y. & Kung, C. *Science* **218**, 153-156 (1982).

Amiloride-sensitive epithelial Na^+ channels reconstituted into planar lipid bilayer membranes

Sarah Sariban-Sohraby*†, Ramon Latorre†, Maurice Burg*, Lori Olans‡ & Dale Benos‡

* Laboratory of Kidney and Electrolyte Metabolism, National Institutes of Health, Bethesda, Maryland 20205, USA

† Departamento de Biología, Facultad Ciencias Básicas y Farmaceuticas, Universidad de Chile, Santiago, Chile

‡ Department of Physiology and Biophysics, Harvard Medical School, Boston, Massachusetts 02115, USA

High resistance epithelia actively transport sodium from the luminal side to the blood¹. Aldosterone and vasopressin stimulate this sodium transport system; the diuretic drug amiloride inhibits it in a reversible fashion. The first step in the transepithelial transport of Na^+ is the facilitated diffusion of Na^+ across the apical membrane via Na^+ -specific, amiloride-sensitive channels²⁻⁴. We report here the first direct measurements of single, amiloride-sensitive Na^+ channel activity. The channel was isolated after incorporation of purified apical membrane vesicles from A6 cells into planar lipid bilayers. The channel had the following characteristics: single-channel conductance ranged from 4 to 80 pS at 200 mM NaCl; it was perfectly cation-selective; amiloride reduced the open-state conductance in a dose-dependent fashion when present in the *cis* compartment, and induced flickering when present in the *trans* chamber; channel conductance and gating were voltage-independent; and the Na^+/K^+ selectivity ratio of the channel was 2:1.

We grew the A6 epithelial cells on Millipore filters because the cells express amiloride-sensitive, apical sodium influx when grown on filters, but not when grown on conventional plastic dishes⁵. Vesicle preparation is summarized in Fig. 1 legend. A6 vesicles were enriched in apical membranes, as evidenced by a 10-fold enrichment in alkaline phosphatase activity without

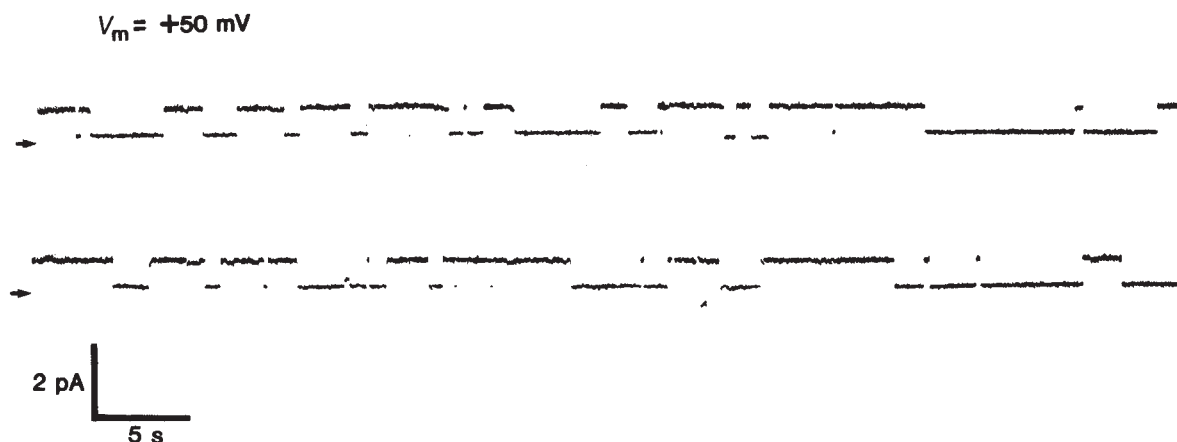


Fig. 1 Cultured A6 apical membrane sodium channels incorporated into planar lipid bilayers. Bilayers were formed from a mixture of bovine brain phosphatidylethanolamine (PE) and phosphatidylserine (PS) at a PE/PS ratio of 70:30 in decane (20 mg lipid per ml). Lipid bilayers were formed on polystyrene cups with circular apertures of 200–340 μm diameter. The bilayer separated two aqueous chambers. When filled to the same height, the *cis* chamber contained 3 ml of solution and the other (*trans*) 4.7 ml. Electrical measurements were made with a two-electrode voltage clamp⁸. The *cis* chamber was connected to a function generator which controlled membrane potential. The *trans* chamber was connected to a current-to-voltage converter, keeping this compartment at virtual ground potential. A positive current was defined as net cation flow into the *trans* side. The aqueous solutions were connected to the electronics by 0.1 M NaCl agar bridges in series with Ag/AgCl electrodes. Apical membrane vesicles were prepared by scraping A6 cells into buffer containing 10 mM CaCl_2 to precipitate nuclei. The scraped material was homogenized and the apical membrane fraction was obtained by differential centrifugation. Membranes were stored (1 mg protein ml^{-1}) at -80°C in 0.5 M sucrose, 10 mM Tris-HEPES, pH 7.4. After bilayer formation in symmetrical salt solutions (0.1 M NaCl, 10 mM MOPS-Tris, pH 7.0), an osmotic gradient was established across the membrane by adding 100–300 mM NaCl to the *cis* chamber to facilitate channel incorporation; 20–30 μl of vesicle suspension were added to the *cis* side. The *cis* solution was vigorously stirred until an initial jump in conductance was observed. Additional incorporation was prevented by stopping the stirring and by extensive washing of the *cis* chamber with a 100 mM NaCl, 10 mM MOPS-Tris, pH 7.0 solution. All single-channel records were stored on FM tape for later analysis. By convention, the channel opening is shown as an upward deflection regardless of the polarity of the applied voltage. The conductance of the bare bilayer was <5 pS in all cases. For the recordings shown here, the holding potential was +50 mV in symmetrical 30 mM NaCl; the arrow indicates the zero-current level. This record was low-pass filtered at 100 Hz.

detectable succinate dehydrogenase activity. Also, after ^{125}I -iodination of apical membrane proteins of confluent A6 monolayers, ^{125}I -specific activity was 7.1 times enriched in the vesicle fraction. Previous evidence for sodium channel activity in A6 vesicles was that amiloride inhibited the ^{22}Na influx with a half-maximal inhibition constant (K_i) of 0.07 μM , similar to that measured in intact A6 cells⁵.

Ionic channels present in these membranes were transferred to planar membranes as described previously^{6–8}. Vesicles were added to the *cis* compartment. The time which elapsed between the addition of vesicles and the first conductance jump (indicating the incorporation of channels into the bilayer) ranged from 10 min to 2–3 h. Channels appeared to incorporate either in bursts (2–6 channels) or one by one. Addition of calcium was not required, and channel incorporation was irreversible.

Figure 1 shows a record of single-channel activity. The average single-channel conductance observed in this experiment was 15 pS in symmetrical 30 mM NaCl, and 55 pS in 200 mM NaCl. The range of single-channel conductances was 4–80 pS at 200 mM NaCl with a mean of 35 ± 8 ($\pm$ s.e.) pS (22 membranes, 8 vesicle preparations). The variability in conductance could be related to the possible existence of a wide spectrum of stable conductance states, as suggested for the gramicidin A channel⁹. Both the conductance of any single channel and its opening probability were independent of voltage (± 60 mV). All channels had similarly long mean open times (2–10 s), absolute cation selectivity, and amiloride sensitivity. At all sodium concentrations, two processes having different kinetics were observed. The channels fluctuated both rapidly and slowly between the open and closed state. The mean residence time of the long-lived closed state was 5.0 ± 1.5 s and was independent of voltage between +30 and –30 mV. The mean closed time of the fast state was <200 ms.

Figure 2a shows mean open single-channel current versus applied potential. Such experiments yield information on

channel cation/anion selectivity. Single open-channel current was a linear function of potential in the range ± 55 mV. The observed zero current potential of -12 ± 1.2 mV agrees well with the predicted reversal potential of -12.8 mV for a perfectly cation-selective channel in these ionic conditions. Current-voltage curves in physiological conditions yielded a reversal potential of -20 mV, indicative of a Na^+/K^+ permeability ratio of 2:1. Single-channel conductances measured in the presence of either KCl or NaCl on both sides of the membrane yielded a conductance ratio of 2.8 Na^+ over K^+ (Fig. 2b). A difference at high salt concentration between the selectivity measured by conductance and by zero-current potential is characteristic of channels that can accommodate no more than one ion at a time. Theoretically, the permeability ratio equals the conductance ratio only in the limit of zero ion concentration¹⁰.

The ability of the channels to discriminate between Na^+ and K^+ is less than in other epithelia^{11,12}. The Na^+/K^+ selectivity of the Na^+ channel in cultured kidney cells is unknown but amiloride has been shown to interact with 'mixed' Na^+-K^+ channels in the apical membrane of toad skin during the moulting process¹³, and in the skin of larval frogs¹⁴.

Amiloride interacts with this channel whether present in the *cis* or in the *trans* compartment. Figure 3a shows the apparent reduction of single-channel open-state conductance by *cis* amiloride. This effect of amiloride was fully reversible. The channels blocked by *cis* amiloride underwent a smooth reduction in open-state conductance as drug concentration was increased (Fig. 3b). The K_i was 0.1 μM , in good agreement with the value of K_i obtained for $^{22}\text{Na}^+$ influxes into A6 vesicles (0.07 μM) and intact cells (0.05 μM)⁵. *cis* Amiloride did not induce fluctuations in open-channel current. Addition of amiloride to the *trans* chamber produced rapid transitions between open and blocked conductance states. These observations may indicate asymmetry of the channel, presence of multiple amiloride binding sites, or different amiloride accessibilities to a single

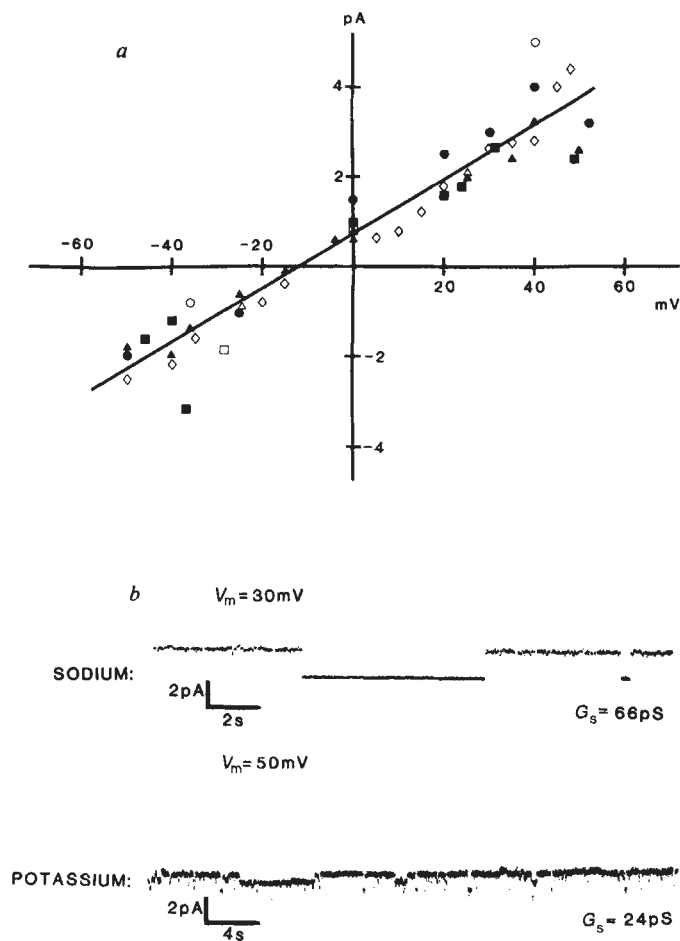


Fig. 2 *a*, Single-channel current-voltage relationship. The different symbols represent data from different membranes. The height of the current jumps for a single channel was measured at different imposed membrane potentials. The NaCl activity gradient was 166:100 mM *cis/trans*. All solutions were buffered to pH 7.0 with 10 mM MOPS-Tris. Linear regression analysis of the data yielded a single-channel conductance of 60 ± 5 ($\pm$ s.e.) pS and a zero-current potential of -12 ± 1.2 mV at a 99% confidence interval. All channels used in this plot displayed reversible amiloride sensitivity. *b*, Single-channel conductance with either Na^+ or K^+ . Channel transitions were recorded in symmetrical solutions of 200 mM Na^+ or K^+ chloride, 10 mM MOPS-Tris, pH 7.0. Channel conductances determined from at least 20 transitions were: Na^+ , 66 ± 2 pS; K^+ , 24 ± 1 pS. Conductance sublevels were not included when calculating mean single-channel conductance.

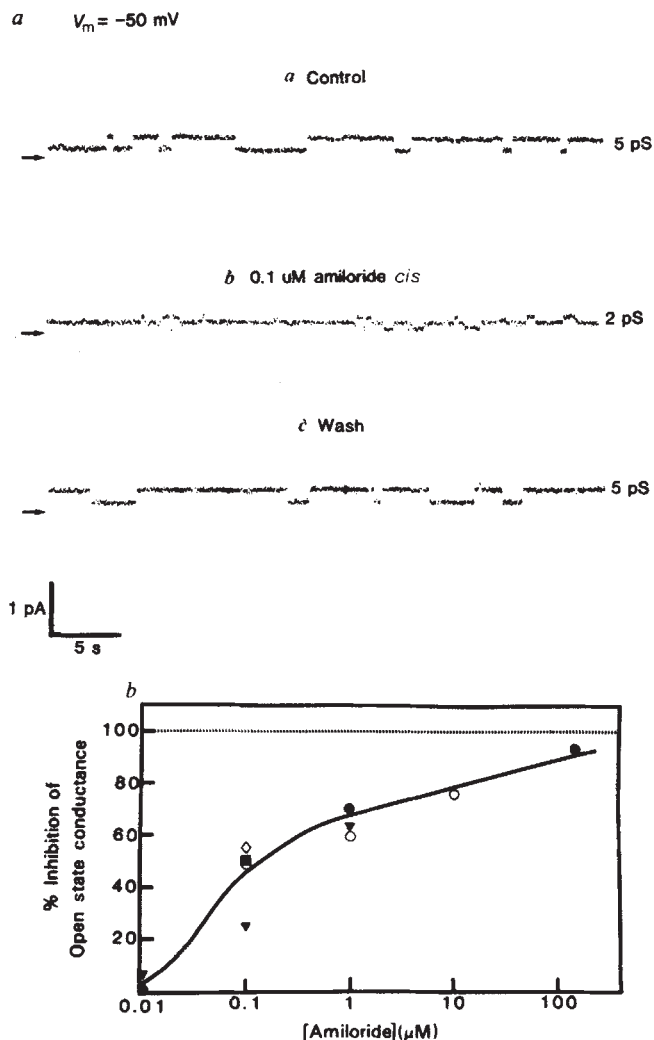


Fig. 3 Amiloride inhibition of single-channel conductance. *a*, Single-channel current fluctuations measured in the absence or presence of amiloride in symmetrical 10 mM NaCl, 10 mM MOPS-Tris, pH 7.0 solutions. This membrane contained two channels (*b*), although only one is seen in *a* and *c*. *a*, Before addition of amiloride. *b*, After addition of amiloride to a final concentration of 0.1 μM to the *cis* side only. *c*, After amiloride was removed by washing the *cis* chamber with 40 ml of drug-free solution. *b*, Per cent inhibition of open-state conductance observed at different concentrations of amiloride with a *cis/trans* 166:100 mM NaCl gradient. Each symbol represents data taken from one of six membranes studied. Amiloride was added to the *cis* side only.

interaction site. The nature of the *cis* blockade by amiloride was unexpected in that the block was much faster (mean blocked state < 1 ms) than suggested by the amiloride-induced current noise in frog skins¹⁵. In our experiments the blocked state is short-lived with respect to the amplifier response time, yielding a mean apparent open-state conductance equal to the time average of true open and blocked conductances^{16,17}. A similar decrease in the time-averaged open-state conductance due to a fast block has been described previously^{10,17}.

Thus, we have shown that direct measurements of amiloride-sensitive sodium channels from A6 apical membranes can be made after incorporation into planar lipid bilayers. The most striking features of this channel are the extremely fast block by *cis* amiloride and the low selectivity of this channel for Na^+ over K^+ . Clearly these observations differ from those obtained indirectly in intact epithelia using current fluctuation analysis.

We thank Drs C. Miller, J. Reyes and D. Shoemaker for helpful comments on the manuscript, and Ms Tamara

Hardenbergh for secretarial assistance. This work was supported by NIH grants GM-25277, GM-28992 and AM-25886, and by funds from the Andrew W. Mellon Foundation.

Received 17 October; accepted 2 December 1983.

1. Macknight, A. D. C., Dibona, D. R. & Leaf, A. *Physiol. Rev.* **60**, 615-715 (1980).
2. Lindemann, B. & Van Driessche, W. *Science* **195**, 292-294 (1977).
3. Van Driessche, W. & Lindemann, B. *Nature* **282**, 519-520 (1979).
4. Benos, D. J. *Am. J. Physiol.* **242**, C131-C145 (1982).
5. Sariban-Sohrab, S., Burg, M. B. & Turner, R. J. *Am. J. Physiol.* **245**, C167-C171 (1983).
6. Miller, C. & Racker, E. *J. Membrane Biol.* **30**, 285-300 (1976).
7. Miller, C. *J. Membrane Biol.* **40**, 1-23 (1978).
8. Latorre, R., Vergara, C. & Hidalgo, C. *Proc. natn. Acad. Sci. U.S.A.* **79**, 805-809 (1982).
9. Busath, D. & Szabo, G. *Nature* **294**, 371-373 (1981).
10. Coronado, R., Rosenberg, R. L. & Miller, C. *J. gen. Physiol.* **76**, 425-446 (1980).
11. Benos, D. J., Mandel, L. J. & Simon, S. A. *J. gen. Physiol.* **76**, 233-247 (1980).
12. Palmer, L. J. *J. Membrane Biol.* **67**, 91-98 (1982).
13. Katz, U. *J. Membrane Biol.* **38**, 1-9 (1978).
14. Hillyard, S. D., Zeiske, W. & Van Driessche, W. *Pflügers Arch. ges. Physiol.* **394**, 287-293 (1982).
15. Lindemann, B. *J. Membrane Biol.* **54**, 1-11 (1980).
16. Neher, C. & Steinbach, J. H. *J. Physiol., Lond.* **277**, 153-176 (1978).
17. Coronado, R. & Miller, C. *Nature* **288**, 495-497 (1980).

Visual pigment in fish iridocytes

J. N. Lythgoe, Julia Shand & R. G. Foster

Zoology Department, University of Bristol, Bristol BS8 1UG, UK

The iridophores of some fishes including the neon tetra, *Paracheirodon innesi*, contain regular alternating layers of guanine and cytoplasm whose spacing changes in response to light¹⁻³, even in decapitated animals³. As a result, the wavelength of light that is most strongly reflected also changes and the iridophore changes colour¹⁻⁴. These iridophores give the colour to the bright iridescent lateral stripe that includes the iris and runs laterally along the body. The iridophores themselves are located between the dermal collagen and the underlying muscle. The photosensitive site appears to be located within the cell itself, or very close to it¹. Using an immunofluorescence technique, we describe here the identification of an opsin-based visual pigment within the neon tetra iridophore.

The antigen used for immunofluorescence was a solution of rhodopsin prepared for us by Dr A. Knowles as follows. Pellets of quail rod outer segments were washed, dissolved in 50 mM Tris buffer at pH 7.0, treated with Ammonyx to break up the membranes, and separated on a concanavalin A (Con A)-Sephadex column⁵. Rabbits were immunized against the antigen according to the method of Lynch and Shirley⁶. We also used rabbit anti-bovine rhodopsin, given by Dr E. Kean⁷, which was prepared by injecting into rabbits bovine rhodopsin, purified by calcium phosphate Celite and Con A-Sephadex chromatography. The immune serum was partially purified by passage over a DEAE-cellulose column. The immunoglobulin fraction was then dialysed against water and lyophilized.

Neon tetras were killed by decapitation and 1-3 mm pieces of retina and skin, taken from the region of the lateral stripe, were incubated in 20 mM ammonium acetate at 4 °C for 15 min to swell the photoreceptor membranes⁸. The tissue blocks were then fixed in 3% glutaraldehyde in isotonic phosphate-phosphate buffer at 4 °C for 18 h, dehydrated in alcohol and embedded in the hydrophilic acrylic resin LR White (London Resin Co.), which seems particularly suitable for immunohistochemistry⁹. Sections (1-2 µm) were placed on glass slides using 0.01% poly-L-lysine (Sigma) as an adhesive¹⁰. The sec-

tions were treated with goat serum diluted 1:40 in phosphate-buffered saline (PBS) (Miles Yeda), rinsed in PBS and incubated in the primary antiserum raised in rabbit. Powdered bovine rhodopsin primary antiserum (10 mg) was dissolved in PBS to give a dilution of 1:4,000 (w/v). The quail rhodopsin primary antiserum was active at dilutions of 1:64 but gave best results at a dilution of 1:32. Incubation was at 37 °C for 30 min. Two controls were used: (1) normal rabbit serum (Miles Yeda) was substituted for the rhodopsin antiserum at the same dilution and in the same incubation conditions; (2) to check that the antiserum was acting specifically against rhodopsin, 1 ml of 1:4,000 (w/v) bovine rhodopsin antiserum was incubated with 1 ml of bovine rhodopsin (prepared in the same way as the quail rhodopsin) at 37 °C for 30 min, centrifuged, and the supernatant used instead of the active antiserum. All slides were thoroughly rinsed in PBS then incubated in fluorescein isothiocyanate (FITC)-conjugated anti-rabbit immunoglobulin raised in goat (Miles Yeda) and diluted at 1:16 in PBS, incubated at 37 °C for 60 min, again rinsed in PBS, mounted in 1 part PBS to 3 parts glycerine and examined under an Olympus fluorescence microscope.

The retina preparations that had been incubated with either anti-bovine or anti-quail rhodopsin antiserum showed strong green fluorescence of all the rod and cone outer segments. No fluorescence was visible in the controls incubated with normal rabbit serum, and there was only weak fluorescence in the preparations in which the anti-rhodopsin antibody had been incubated with rhodopsin. This makes it unlikely that the antibody was active against molecules other than opsin in either the iridophores or the outer segments. In the neon tetra both the melanophores and the iridophores lie between the dermal collagen and the underlying muscle³. Sections treated with anti-rhodopsin showed a double rank of strongly green fluorescent cells having the same shape and in the same position as the iridocytes (Fig. 1a, b). The green fluorescence extended across the entire central area of the cell and showed a diagonal banded pattern (Fig. 1c) characteristic of the arrangement of the iridescent multilayers viewed under the electron microscope (Fig. 2). The individual reflecting plates are less than 100 nm thick^{3,4}; thus they are too small to be resolved under the light microscope. The diagonal fluorescent bands seen in Fig. 1c probably represent several plates that have remained grouped together.

Although the antibody was raised against rhodopsin, it seemed to be active against all the cones in the retina, and these contain

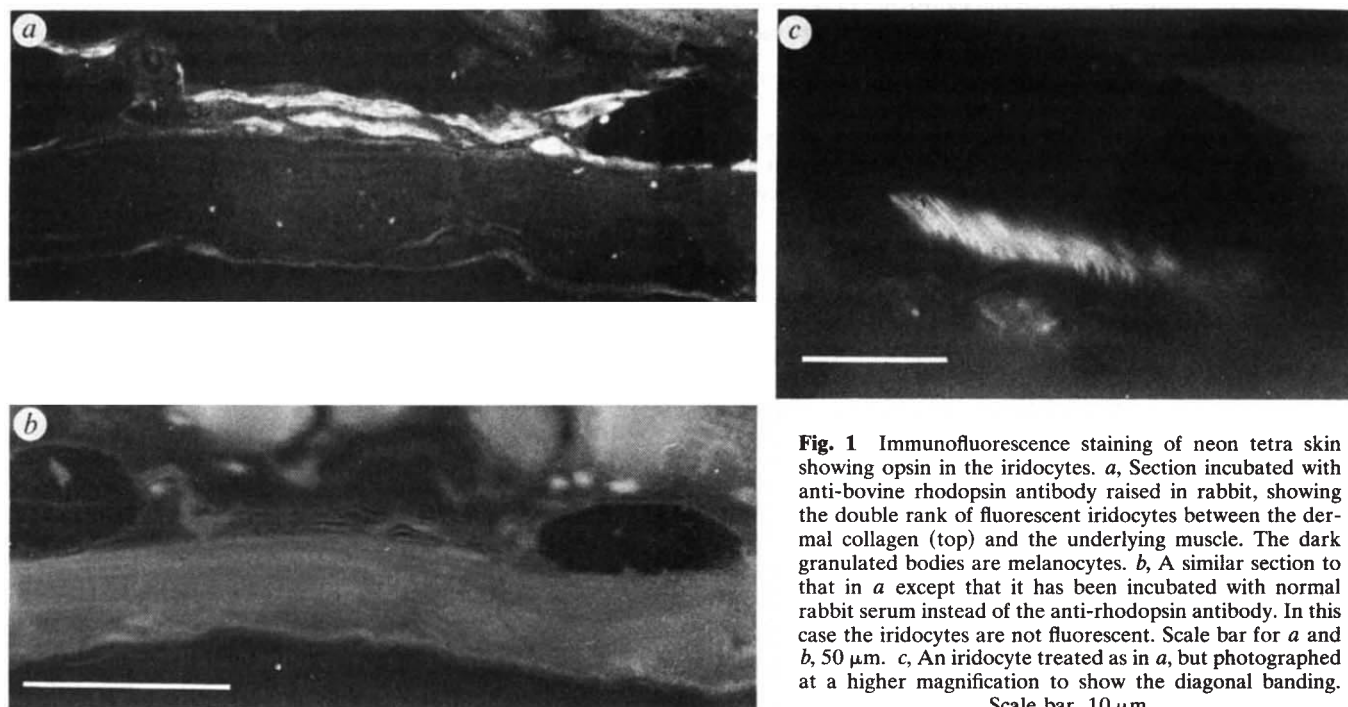


Fig. 1 Immunofluorescence staining of neon tetra skin showing opsin in the iridocytes. *a*, Section incubated with anti-bovine rhodopsin antibody raised in rabbit, showing the double rank of fluorescent iridocytes between the dermal collagen (top) and the underlying muscle. The dark granulated bodies are melanocytes. *b*, A similar section to that in *a* except that it has been incubated with normal rabbit serum instead of the anti-rhodopsin antibody. In this case the iridocytes are not fluorescent. Scale bar for *a* and *b*, 50 µm. *c*, An iridocyte treated as in *a*, but photographed at a higher magnification to show the diagonal banding. Scale bar, 10 µm.



Fig. 2 Electron micrograph section of an iridocyte showing the reflecting multilayer stack. A possible site for the visual pigment is in the membranes within the reflecting stack. Scale bar, 500 nm.

A second plasma calcium-lowering peptide from the human calcitonin precursor—a re-evaluation

I. MacIntyre, C. J. Hillyard, J. J. Reynolds, R. E. Gaines Das & R. K. Craig

Nature **300**, 460–462 (1982)

WE recently reported that the C-terminal flanking peptide (katakalcin or PDN-21) from the human calcitonin precursor circulates in man in equimolar relation to calcitonin and that the plasma katakalcin level rises after calcium infusion and is grossly elevated in medullary carcinoma of the thyroid (*Lancet* **i**, 846–848, 1983). These findings are in good agreement with our earlier report (reference above). In that paper, however, we also reported that katakalcin lowered plasma calcium after intravenous injection in 50 g rats and described a direct effect of the synthetic peptide on mouse calvariae in organ culture, although at much larger concentrations.

We must now report that in subsequent experiments using different preparations of synthetic katakalcin, the calcium lowering activity has been highly variable at best. Additionally we do not find any effect on bone resorption using mouse calvariae in organ culture. Whatever the reasons for the differences between our original studies and the present ones, we now feel it is unsafe to extrapolate from our results to a physiological hypocalcaemic effect of katakalcin in man. □

Errata

Acceptors for botulinum neurotoxin reside on motor nerve terminals and mediate its internalization

J. O. Dolly, J. Black, R. S. Williams & J. Melling

Nature **307**, 457–460 (1984)

THE abbreviation for botulinum toxin in line 1 of the bold first paragraph should read BoNT, and in the legend to Fig. 1 the line 'a, c, e, With tetanus toxin; b, d, f, without.' should be deleted.

Inverse relationship between neurotensin receptors and neurotensin-like immunoreactivity in cat striatum

M. Goedert, P. W. Mantyh, P. C. Emson & S. P. Hunt

Nature **307**, 543–546 (1984)

IN the sixth line of the 'Methods' section of Fig. 1 legend, the sentence 'Control experiments had established that there was no detectable neurotensin- or (Met)enkephalin-like sections' should read 'Control experiments had established that there was no detectable neurotensin- or (Met)enkephalin-like immunoreactivity left in the tissue sections'.

Cloning and sequence analysis of calf cDNA and human genomic DNA encoding α -subunit precursor of muscle acetylcholine receptor

M. Noda *et al.*

Nature **305**, 818–823 (1983)

THE sentence on lines 20–22 of Fig. 4 legend should read 'The degree of amino acid sequence homology in each of these protein regions is given below for the human/calf (dotted lines), human/*Torpedo* (solid lines) and calf/*Torpedo* (dashed lines or solid lines where no dashed lines are shown) pairs; . . . '.

at least some porphyropsin rather than rhodopsin¹¹. We believe, therefore, that we have demonstrated the presence of an opsin-based photopigment in the iridocyte, but it is unknown whether it is rhodopsin or porphyropsin.

Birds and mammals also possess dermal photoreceptors but the mechanism of the light response has not been extensively studied^{12–15}. Light induces electrical changes in frog skin¹⁶, and action spectra of the electrodermograms have prompted the suggestion that a blue-sensitive rhodopsin and possibly a rhodopsin–metarhodopsin photochromic system is involved¹⁷. The direct action of light on the melanophores of larval *Xenopus*^{18,19} and of at least four species of larval frogs, causes the melanosomes to aggregate within the melanocyte and the skin colour to lighten^{18,19}. These photosensitive sites are localized within the cell itself^{20,21}, but the nature of the photoreceptive substance is unknown. Perhaps these melanocytes, like the neon tetra iridocytes, also contain visual pigment.

We thank Dr E. L. Kean for a sample of anti-bovine rhodopsin antiserum, Dr A. Knowles for purified quail and bovine rhodopsin; and Professor B. K. Follett and Dr A. E. Dorey for help and advice. This research was supported by MRC project grant G8125739 (J.N.L. and J.S.) and by an ARC research studentship (R.G.F.).

Received 30 September; accepted 25 November 1983.

1. Foster, K. W. *J. exp. Zool.* **77**, 169–214 (1937).
2. Rohlich, S. T. *J. Cell Biol.* **62**, 295–304 (1974).
3. Lythgoe, J. N. & Shand, J. *J. Physiol., Lond.* **325**, 23–34 (1982).
4. Land, M. F. *Prog. Biophys. molec. Biol.* **24**, 75–106 (1972).
5. de Grip, W. J. *Meth. Enzym.* **81**, 197–207 (1982).
6. Lynch, S. S. & Shirley, A. *J. Endocr.* **65**, 127–132 (1975).
7. Lentricchia, B. B., Plantner, J. J. & Kean, E. L. *Expl Eye Res.* **31**, 1–8 (1980).
8. Rohlich, P. *Nature* **263**, 789–791 (1976).
9. Newman, G. R., Jasani, B. & Williams, E. D. *Histochem. J.* **15**, 543–555 (1983).
10. Husain, O. A. N., Millett, J. A. & Granger, J. M. *J. clin. Path.* **33**, 309–311 (1980).
11. Levine, J. S. & MacNichol, E. F. *Jr Sensory Process* **3**, 95–130 (1979).
12. Fingerman, M. A. *Rev. Physiol.* **32**, 345–372 (1970).
13. Adler, K. *Photochem. Photobiol.* **23**, 275–298 (1976).
14. Wolken, J. J. & Magus, M. A. *Photochem. Photobiol.* **29**, 189–196 (1979).
15. Steven, D. M. *Biol. Rev.* **38**, 204–240 (1963).
16. Becker, H. E. & Cone, R. A. *Science* **154**, 1051–1053 (1966).
17. Wald, G. & Rayport, S. *Biol. Bull.* **147**, 503 (1974).
18. Bagnara, J. T., Taylor, J. D. & Protta, G. *Science* **182**, 1034–1035 (1973).
19. Bagnara, J. T. *J. exp. Zool.* **187**, 149–154 (1974).
20. van der Lek, B., de Heer, J., Burgers, A. C. J. & van Oordt, C. J. *Acta physiol. pharmac. Neerl.* **7**, 409–419 (1958).
21. Bagnara, J. T. & Obika, M. *Experientia* **23**, 155–157 (1967).

No real homology of human T-cell leukaemia virus envelope with class I HLA

In a recent paper Clarke *et al.*¹ presented hybridization data purporting to show sequence homology between a portion of a human T-cell leukaemia virus (HTLV) and a human class I HLA gene (HLA-B7). We have directly compared the published sequences of the appropriate portions of HTLV² and class I HLA nucleotide and/or protein sequences (HLA-B7 and HLA-A2—whose protein sequence is very closely related to HLA-B7³) using the computer programs DIAGON⁴ and ANALYSEQ⁵. We find no significant homology of nucleotide sequence between HTLV and HLA-A2, nor can we find any protein sequence homology between HTLV and HLA-A2 or HLA-B7. As we do not have the precise nucleotide sequence of the HLA-B7 clone, we cannot formally exclude the possibility of the DNA sequence homology claimed by Clarke *et al.*¹. However, the lack of protein sequence homology makes it most unlikely that the *env* gene of HTLV is structurally related to a histocompatibility antigen. We therefore wonder why no sequence comparison was presented by the authors or requested by *Nature*. It is now common knowledge that G+C-rich sequences can confuse cross-species hybridization results, as is admitted by Clarke *et al.*¹. Indeed we note that there are G+C-rich sequences in the hybridizing fragments of HTLV described by Clarke *et al.*¹. For example, *XhoI*-*XhoI*, 901–1,000, T=19%, C=36%, G=36%, A=9%; *XhoI*-*SstII*, 1,800–1,930, T=31.3%, C=40.5%, G=22.9%, A=5.3% (nucleotide numbering starting at the A of the initiation codon of the envelope gene²). Clearly, hybridization at extremely low stringency is not a sufficient criterion for statements of sequence homology. It can, however, serve as a useful preliminary to cloning and thereby obtaining the necessary sequence data.

P. A. JENNINGS
J. EVEN

Laboratory of Molecular Biology,
Medical Research Council Centre,
Hills Road, Cambridge CB2 2QH, UK

1. Clarke, M. F., Gelmann, E. P. & Reitz, M. S. Jr *Nature* **305**, 60–62 (1983).
2. Seiki, M., Hattori, S., Hirayama, Y. & Yoshida, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3618–3622 (1983).
3. Malissen, M., Malissen, B. & Jordan, B. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 893–897 (1982).
4. Staden, R. *Nucleic Acids Res.* **10**, 2951–2961 (1982).
5. Staden, R. *Nucleic Acids Res.* (in the press).

CLARKE, GELMANN AND REITZ JR
REPLY—At the time our manuscript¹ was accepted for publication, no sequence data were available for human T-cell leukaemia virus (HTLV). By the time we received proofs the sequence data of Seiki *et al.*²

had been published. We compared the sequence of the *env* region of HTLV with the published sequence data for the HLA-B7 pseudogene clone of Malissen *et al.*³, did not find what we felt to be sufficient homology to explain the hybridization data, and consequently appended a *Note added in proof* which, however, was not published, at the discretion of *Nature*. There are several likely reasons for the apparent discrepancy between hybridization and sequence data. First, the published sequences are not of the clones used in our studies. Our probe is one for HLA-B7 (ref. 4) while the clone of Malissen *et al.* represents a pseudogene with significant amino acid sequence differences from HLA-B7 (ref. 3). The clone of Seiki *et al.*² was not tested by us, differs in the region of homology by at least three restriction enzyme sites, and may lack the homology we have described. Indeed, a clone⁵ of HTLV-II⁶, which does not alter the HLA class I expression in infected cells (our unpublished data with D. Mann, M. Popovic and R. C. Gallo), does not hybridize with our HLA probe. Second, relevant to a possible structural relationship between an HTLV *env* protein and HLA class I antigens, it should be pointed out that B7 is not one of the 'extra' class I specificities detected on HTLV-I-infected cells⁷, and, were the sequences of the appropriate antigens available, a greater amino acid sequence homology would probably be evident. Moreover, the stretch of homology need only be sufficient to account for a single antigenic site.

As to the specificity of hybridization, contrary to the assertions of Jennings and Even, hybrid formation was performed at moderate stringency (4×SSC, 50% formamide, 37°C) and washing at moderately low stringency (2×SSC, 50°C), and was therefore not at 'extremely low stringency'. We feel that we have ruled out the possibility of G+C-rich nonspecific hybridization in any event by demonstrating a high degree of specificity. As mentioned above, no hybridization is observed with the related HTLV-II. Hybridization to the *XhoI*-*XhoI* region (901–1,000) cited by Jennings and Even as unusually G+C-rich is barely detectable (see ref. 1, Fig. 1, lane c). There are areas elsewhere in the HTLV-I genome as G+C-rich as the areas indicated by Jennings and Even; for example, bases 758–901 of Seiki *et al.*² (64% G+C), bases 1,070–1,198 (69% G+C and bases 1,530–1,630 (68%), all in the *gag* region and none hybridizing with the probe. No hybridization is obtained with clones for DR- α ⁸ and DC-B⁹, the latter being as G+C-rich as the HLA-B7 probe. While we realize (and stated in our paper) that even specific hybridization cannot prove that the two proteins in question are partially structurally homologous, we feel that these data (together with those of Mann *et al.*⁷) strongly suggest that this is indeed the case. We are currently placing appropriate

portions of the HTLV genome into mammalian expression vectors in order to resolve this question definitively.

M. F. CLARKE
E. P. GELMANN
M. S. REITZ JR

Laboratory of Tumor Cell Biology,
National Cancer Institute,
National Institutes of Health,
Bethesda, Maryland 20205, USA

1. Clarke, M. F., Gelmann, E. P. & Reitz, M. S. Jr *Nature* **305**, 60–62 (1983).
2. Seiki, M., Hattori, S., Hirayama, Y. & Yoshida, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3618–3622 (1983).
3. Malissen, M., Malissen, B. & Jordan, B. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 893–897 (1982).
4. Sood, A. K., Pereira, D. & Weissman, S. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 616–620 (1981).
5. Gelmann, E., Franchini, S., Manzdi, V., Wong-Staal, F. & Gallo, R. C. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
6. Kalyanaraman, V. S. *et al. Science* **218**, 571–573 (1983).
7. Mann, D. *et al. Nature* **305**, 58–60 (1983).
8. Gustafson, K. *et al. Scand. J. Immun.* **16**, 303–308 (1982).
9. Larhammar, D. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3677–3691 (1982).

No role of vasopressin in stress-induced ACTH secretion?

RECENTLY, Mormède¹ reported that exposure of conscious rats to a novel environment induced the secretion of adrenocorticotrophic hormone (ACTH) and corticosterone. As this secretion could not be blocked by an anti-pressor vasopressin antagonist, 1-deaminopencillamine, 2-(O-methyl)tyrosine arginine-vasopressin (dPyr(Me)AVP), the author concluded that vasopressin was not physiologically involved in ACTH release. We disagree with this contention: as shown below, pituitary vasopressin receptors are not effectively blocked by anti-pressor vasopressin antagonists.

Experiments were performed on dispersed anterior pituitary cells² (Geneva), on hemipituitaries from 6–7-day-old rats cultured for 4 weeks *in vitro*³ (Basel), and on anterior pituitary segments⁴ (Budapest). The tissues and cells were incubated for 1 h or 2 h in the presence of 2.5–20 nM arginine vasopressin (AVP; Sigma), 0.5–20 nM synthetic ovine corticotropin-releasing factor (CRF; Bachem), 50–200 nM anti-pressor AVP antagonists^{5,6}, and combinations thereof. All incubations were in 10–100 μ M ascorbic acid to prevent oxidation and inactivation of CRF⁷. The antagonist dPyr(Me)AVP was synthesized by Dr J. Rivier (San Diego) and the antagonist d(CH₂)₅Tyr(Me)AVP by Dr Manning (Toledo). After the incubation period, the media were frozen (–20°C) pending radioimmunoassay for ACTH (994 antiserum from Dr C. Oliver, Marseille, for the Basel and Geneva experiments, and antiserum no. 6 for the Budapest experiments).

Table 1 Release of ACTH from dispersed anterior pituitary cells

Experimental condition	Release of ACTH (pg ml ⁻¹)
AA alone	1,504 ± 164
AVP+AA	2,727 ± 199
AP+antagonist+AA	2,430 ± 156
CRF+AA	3,461 ± 167
CRF+AVP+AA	7,295 ± 339
CRF+AVP +antagonist+AA	6,339 ± 522

Anterior pituitary cells were incubated for 1 h with 5 nM AVP, 5 nM CRF, 200 nM d(CH₂)₅Tyr(Me)AVP (antagonist) and combinations thereof in the presence of 10 μM ascorbic acid (AA). Results are mean ± s.e.m. for four separate cell suspensions each.

A representative experiment on dispersed anterior pituitary cells is shown in Table 1. The anti-pressor AVP antagonist d(CH₂)₅Tyr(Me)AVP had no effect on ACTH release evoked by AVP. Similar results were also obtained with the dPTyr(Me)AVP antagonist in all three *in vitro* systems. The potentiation of the ACTH releasing effect of CRF by AVP was slightly and not significantly reduced by both antagonists in dispersed anterior pituitary cells and pituitary fragments, and was reduced by 20–40% ($P < 0.05$) in cultured hemipituitaries.

Earlier studies *in vivo* had shown that anti-pressor antagonists administered in 5–50-fold excess of AVP blocked the stimulation of ACTH and corticosterone secretion by AVP^{1,8,9}. It was suggested that the site of the antagonism was at the pituitary level. The present *in vitro* studies, however, indicate that these antagonists are unlikely to abolish the direct effect of AVP on ACTH release either in the presence or absence of CRF. In fact, previous studies have also pointed out that pituitary vasopressin receptors are neither of the pressor (V₁) nor of the antidiuretic (V₂) type^{10–12}, and antidiuretic agonists have less intrinsic or potentiating activity than AVP (our unpublished *in vitro* studies).

We suggest that exogenous AVP and anti-pressor AVP antagonists administered *in vivo* act on ACTH release at a peripheral or central nervous site. Thus the main argument for Mormède's contention¹ cannot be accepted, and two other statements of his are debatable: first, the absence of AVP secretion into peripheral blood may co-exist with the presence of AVP release into hypophysial portal vessels and with hypophysial-adrenal activation¹³. Second, many studies suggest a role for AVP in ACTH secretion *in vivo* (see references in ref. 14).

This study was supported in part by SNSF grant 3.898.81. M.C.H. is an exchange Fellow of the Royal Society (London) and the Hungarian Academy of Sciences. We thank Ms M. Friedli (Geneva), L. Rietschin (Basel) and V.

Garamvölgyi (Budapest) for technical assistance.

A. J. BAERTSCHI

Department of Animal Biology,
University of Geneva,
1211 Geneva 4, Switzerland

B. GÄHWILER

Preclinical Research,
Sandoz A. G., 4002 Basel,
Switzerland

F. A. ANTONI
M. C. HOLMES
G. B. MAKARA

Hungarian Academy of Sciences,
Institute of Experimental
Medicine,
Budapest, Hungary

1. Mormède, P. *Nature* **302**, 345–346 (1983).
2. Bény, J.-L. & Baertschi, A. *J. Neuroendocrinology* **20**, 108–112 (1980).
3. Gähwiler, B. *J. Neurosci. Meth.* **4**, 329–342 (1981).
4. Antoni, F. A., Holmes, M. C. & Jones, M. T. *Peptides* (in the press).
5. Bankowski, K., Manning, M., Haldar, J. & Sawyer, W. H. *J. med. Chem.* **21**, 850–853 (1978).
6. Kruszynski, M. et al. *J. med. Chem.* **23**, 364–368 (1980).
7. Vale, W., Spiess, J., Rivier, C. & Rivier, J. *Science* **213**, 1394–1397 (1981).
8. Aizawa, T., Yasuda, N., Greer, M. A. & Sawyer, W. H. *Endocrinology* **110**, 98–104 (1982).
9. Knebel, W., Benner, K. & Heriting, G. *Eur. J. Pharmac.* **81**, 645–646 (1982).
10. Arimura, A., Schally, A. V. & Bowers, C. Y. *Endocrinology* **84**, 579–583 (1969).
11. Doepfner, W., Stürmer, E. & Berde, B. *Endocrinology* **72**, 897–902 (1962).
12. Knebel, W., Homolka, L. & Vlasovska, M. *Eur. J. Pharmac.* **91**, 115–118 (1983).
13. Burlet, A. et al. in *Neuroendocrinology of Vasopressin, Corticotropin and Opiomelanocortin* (eds Baertschi, A. J. & Dreifuss, J. J.) 95–105 (Academic, New York, 1982).
14. McCann, S. M. *Neuroendocrinology* **31**, 355–363 (1980).

MORMÈDE REPLIES—Baertschi and collaborators report here that antagonists of pressor-like receptors of vasopressin (AVP) are unable to block ACTH release and CRF potentiation induced by AVP *in vitro*. Consequently, the inability of such an antagonist to modify the pituitary-adrenocortical response to a psychological stressor¹ cannot be taken as an argument for the non-involvement of pituitary AVP receptors in ACTH release. However, this issue was considered in the conclusion of my report¹: “the inability of the antagonist to change the ACTH response to the psychological stress of exposure to a novel environment can be explained by the non-involvement of hypophysial AVP receptors in ACTH release... On the other hand, the hypophysial AVP receptor involved in CRF potentiation could be insensitive to the antagonistic properties of dPTyr(Me)AVP” and I have recently gathered evidence to support this later possibility in conscious, freely moving animals.

AVP has by itself a low ACTH-releasing activity *in vitro* but potentiates CRF-induced ACTH release². However, the characteristics of the pituitary vasopressin receptor have so far been poorly defined. Two subtypes of AVP receptors have been shown to mediate the peripheral effects of vasopressin (see review in ref. 3). Analogues having differential agonistic⁴ or antagonistic^{5,6} activity towards both receptors have been synthesized.

The effect of AVP on ACTH release was classically considered to be mediated by a pressor-like receptor at the anterior pituitary level⁷. However, recent data have challenged this view: the release of ACTH by AVP can be antagonized by either anti-pressor antagonists¹ or by immunoneutralization of CRF with a specific antiserum raised in rabbits against synthetic ovine CRF⁸, which demonstrates that the main effect of AVP on ACTH release is mediated by CRF liberation via a pressor-like receptor. Furthermore, I have recently obtained experimental evidence (P.M., M. LeMoal and R. Dantzer, in preparation) supporting an action of AVP agonists with low pressor activity (such as dDAVP) on ACTH secretion independently of CRF release and of pressor-like receptors. The same analogue of AVP was recently shown to potentiate CRF-induced ACTH release *in vitro*⁹. However, we also found that an antagonist of the antidiuretic-like receptor, d(CH₂)₅-DLeu-VAVP⁶, was unable to antagonize the effect of dDAVP on ACTH release. These results favour the hypothesis of a direct action of vasopressin and its analogues on the pituitary through a receptor having characteristics different from those of the two vasopressin receptors so far defined.

The data obtained *in vitro* by Baertschi and collaborators agree with these findings. However, further research is obviously necessary to better characterize the pituitary AVP receptor in order to reinvestigate the physiological role of vasopressin in stress-induced ACTH secretion.

P. MORMÈDE

Neurobiologie des
Comportements Adaptatifs,
U259 INSERM-INRA,
Rue Camille Saint-Saens,
33076 Bordeaux, France

1. Mormède, P. *Nature* **302**, 345–346 (1983).
2. Gillies, G. E., Linton, E. A. & Lowry, P. J. *Nature* **299**, 355–357 (1982).
3. Jard, S. J. *Physiol., Paris* **77**, 621–628 (1981).
4. Sawyer, W. H., Acosta, M. & Manning, M. *Endocrinology* **95**, 140–149 (1974).
5. Bankowski, K., Manning, M., Haldar, J. & Sawyer, W. H. *J. med. Chem.* **21**, 850–853 (1978).
6. Manning, M., Klis, W. A., Olma, A., Seto, J. & Sawyer, W. H. *J. med. Chem.* **25**, 414–419 (1982).
7. Aizawa, T., Yasuda, N., Greer, M. A. & Sawyer, W. H. *Endocrinology* **110**, 98–104 (1982).
8. Rivier, C. & Vale, W. *Neuroendocrinology* (in the press).
9. Lutz-Bucher, B. & Koch, B. *Neuroendocr. Lett.* **5**, 111–115 (1983).

Game theory in theology

H.N.V. Temperley

Superior Beings: If They Exist, How Would We Know? By Steven J. Brams.
Springer-Verlag: 1983. Pp.202. Pbk DM 32, \$12.50.

MANY scientists hold that no problem is so obscure, or so sacred, that it can be exempt from fearless rational enquiry. Game theory, created by such profound thinkers as von Neumann and Turing in the late 1940s, attempts to codify, and to quantify where appropriate, the many ways in which two or more people or organisms can interact with one another. In this book, Steven Brams has applied its tenets to exploring the ramifications of the possible existence of "superior beings"; while I am not shocked by the title of his book, I am surprised that nothing of this kind seems to have been attempted before.

Many writers, myself included, have tried to look at the problems that would confront a "superior being" (Brams's expression) who was directing an organization as big as the Earth. It may well be intrinsically impossible for the superior being to allow for some free-will on the part of the inhabitants and at the same time ensure a fair distribution of rewards and punishments. And it may well be that the world is learning by its mistakes. Along such lines one can look for a partial solution of the problem of Evil.

The author confines himself almost entirely to the simplest possible type of "game" between an ordinary person (P) and a superior being (SB), in which, at each play, each participant has only two possible courses of action available. Even under these severe restrictions, which obviously preclude consideration of some very important issues such as redemption, priesthood and the possible existence of the Devil, Brams is able to set up and consider a wide variety of intriguing questions, such as how could SB best "reveal" himself to P and what policy He should follow if P commits "sin".

The principal difficulty with this approach is that scientists genuinely do not know what theologians actually mean by concepts such as omniscience and omnipotence. The author examines the effect of using various alternative definitions of his own, always taking great care to ensure that they are consistent with the exercise of free-will by P. Not the least interesting part of the investigation is the study of how changes in the definitions can affect the outcome of some of the games.

The newcomer to game theory will find many surprises in this book. Two participants can play no less than 78 different games under the severe restrictions mentioned above. As many as 57 of these can be described as mathematically non-trivial. In certain games, SB may be obliged

to act capriciously (in game-theory language this is to use a mixed strategy) in order to obtain the best possible outcome. Even more surprising is the conclusion that in certain games the possession by SB of "superior" qualities can lead to a worse outcome for SB, if it is assumed that P knows about, or guesses at, those qualities.

We can discuss the game of "Chicken" as an illustration. This models two players who are on a collision course; the player who first gives way is considered to have lost face. Each player can choose between the two strategies of co-operation (C) and non-cooperation (N). The combined outcome NN is regarded as a disaster by both players who both give it the rating 1 (worst). The outcome CN is rated 2 (next worst) by the first player who chose C but 4 (best) by the second player who chose N. The ratings of NC are mirror images of those of CN, while the outcome CC is rated 3 (second best) by both players. (A game would be trivial, "not worth playing", if both players gave the same rating to all four outcomes.)

This apparently straightforward game gives rise to a paradox. Brams gives a simple argument to show that the likely outcome of play between equals is CC. Now suppose that it is played between SB, who is omniscient in the sense that he can predict P's choice, and that P is aware of this. Now P will undoubtedly choose the strategy N. Why? Because SB, knowing this, will be obliged to choose C to avoid the disastrous outcome NN.

Thus, in this not unrealistic model of a real encounter, SB's "omniscience" has

both damaged his chances and improved those of P, the outcome CN being rated (2,4) in place of the (3,3) obtained if played between equals! This is just one of many fascinating situations analysed by Brams. One can go on to consider what happens if a game is repeated and players are allowed to change their strategies. The conclusion must be that many of the difficulties facing an administrator of the world are encapsulated in a two-person game such as this.

Brams repeatedly refers to an earlier work, *Biblical Games: A Strategic Analysis of Stories in the Old Testament* (MIT Press, 1980), and reports his conclusion that the Jehovah of the Old Testament was "a superlative strategist . . . driven by very humanlike calculations . . .". This earlier study has, apparently, greatly helped Brams in formulating the problems studied in the present book.

A few decades ago a work such as this would have been rejected (probably unread) as frivolous and profane. Had it appeared a little earlier still, the consequences might have been even more drastic. Brams is well aware of this and his introduction, explaining what he is trying to do, is a joy to read. His arguments, some of them quite complicated, are presented clearly and enough background information is given to enable the non-expert in game theory to follow what is going on.

The claim that the book makes a contribution to solving the problem of Evil is justified, but I can hold out no hope that certain aspects of it will ever be dealt with along such lines. Why, for example, did the horrors perpetrated by the Nazis not provoke a drastic divine intervention? And why have the people of Vietnam been punished for such a long time? Game theory cannot begin to answer such formidable questions as these. □

H.N.V. Temperley was formerly Professor of Applied Mathematics at University College, Swansea.

Genetics and ecology

Nick Barton

The Evolutionary Biology of Colonizing Species.

By P.A. Parsons.

Cambridge University Press: 1983.

Pp.262. £15, \$29.95.

How do plants and animals colonize new places, and new habitats? What are the consequences of colonization for the genetic structure of populations, and the ecological structure of communities? These are two questions which touch upon many difficult issues in evolutionary biology. For example, the boundary of a species will often be set by its ability to adapt to extreme environments; given the

rapid and extensive changes which can be effected by artificial selection, the definite borders of many species have seemed surprising, and have led to the idea that although peripheral populations *could*, if left to themselves, adapt and spread into unoccupied territory, the flow of genes from more central populations prevents adaptations to peripheral conditions, and so limits the range of the species. This now seems unlikely; there are many observations of differential adaptation in adjacent populations — for example, the sharp gradients in heavy-metal tolerance in grasses living on mines.

We are left, then, with the problem of understanding what restricts the power of natural selection in individual populations. In some sense, the limit must be set by lack of suitable genetic variation; but to what extent does the primary restriction come

from competition with other species, from simple physical constraints or from both?

Populations do sometimes colonize new areas, with a habitat more or less different from the old one; what are the genetic and ecological consequences? Do new species often form as a result of the drastic changes experienced during a colonization event? Is the continual process of extinction and recolonization more important than simple diffusion in spreading genes through a population (singly or in combination), as envisaged by Wright in his "shifting balance" theory? Are the proportions of different species at an equilibrium resulting from steady competitive pressure, or can community structure only be understood in terms of continual, stochastic turnover?

Given the range of these questions, all of which might be considered under the title *The Evolutionary Biology of Colonizing Species*, this book is disappointingly narrow in scope. Although Parsons sets out to "integrate population genetics and ecology in order to develop a more coherent theory of evolutionary biology", a laudable aim, his book is almost entirely confined to considering the response of a few species, each considered in isolation to the physical environment. It was Darwin's view that "the range of the inhabitants of any country by no means exclusively depends on insensibly changing conditions, but in large part on the presence of other species"; without considering interactions between species, one can hardly understand their ecology.

Nearly all species considered by Parsons are in the genus *Drosophila*. A great deal of information is presented comparing the ranges of different *Drosophila* species in the wild to their responses to physical stresses in the laboratory (alcohol vapour or extremes of temperature, for example). Although this comparative method can be fruitful, we may need a better understanding of the relationship between laboratory and field conditions, in a larger number of species, before the results can be very informative.

A more fundamental criticism of the book is that difficult distinctions, such as those between "hard" and "soft" selection, "structural" and "regulatory" genes, "genic" or "organismic" selection, and "peripheral" and "central" populations, are introduced without any clear explanation of their meaning. This is not only perplexing for a reader unfamiliar with the terminology, but also makes the discussion disturbingly vague; the difficulty of testing many of the central questions against field data makes it especially important to frame such questions unambiguously. Understanding the evolutionary biology of colonizing species will certainly require a synthesis of population genetics and ecology; but that synthesis remains to be achieved. □

Nick Barton is in the Department of Genetics and Biometry at University College London.

Atmospheres around the Solar System

Conway B. Leovy

The Origin and Evolution of Planetary Atmospheres.

By Ann Henderson-Sellers.

Adam Hilger, Bristol/Heyden, Philadelphia: 1983. Pp.236. £19.50, \$34.

WHY is Earth unique among the planetary bodies of the Solar System in providing an environment suitable for life? Why has its climate remained stable over the past three or four billion years despite probable variations in solar output and possible cometary or asteroidal collisions? How likely is it that stars other than the Sun possess habitable zones? Such far-reaching questions have been stimulated by the comparative study of planetary atmospheres and are addressed by Ann Henderson-Sellers in this survey of a complex and highly interdisciplinary field.

All of the planetary bodies in the Solar System are considered, including the major satellites which do not now have atmospheres. Only the outer planets, from Jupiter to Neptune, have primary atmospheres little influenced by such evolutionary processes as surface interactions and escape. Titan is a fascinating special case, with a slowly evolving atmosphere of nitrogen, argon and hydrocarbons.

The questions most relevant to the existence of life centre around the role and

fate of water on the terrestrial planets. Of these, Venus appears to have undergone catastrophic heating at an early stage, leading to evaporative loss of what may have been a rich initial endowment of water. Water has apparently been an important agent of modification of the Martian surface, though geological interpretation of spacecraft images by Michael Carr and others, and surface adsorption models of Fraser Fanale suggest that liquid water may not have been widespread for at least the past three billion years. Nevertheless, large fluctuations in surface pressure and climate may occur even now. Earth's climatic stability is all the more impressive in the context of its neighbours' catastrophic and fluctuating climates, and Henderson-Sellers focuses on the processes which control planetary surface temperatures, arguing that Earth's unique hydrosphere may be primarily responsible for climatic stability.

Those seeking a self-contained overview of the field will be disappointed. In many cases, the explanations of the physical and chemical controls on atmospheric evolution and climate are obscure or incomplete, and one is frequently taken around issues rather than to their core. Nevertheless, the sweeping perspective and interdisciplinary scope of this volume will reward the careful reader by calling to mind a host of intriguing problems worthy of further exploration. □

Conway B. Leovy is in the Department of Atmospheric Sciences at the University of Washington, Seattle.

Hunt for nutrition

Diane Gifford-Gonzalez

Bison Kills and Bone Counts: Decision Making by Ancient Hunters.

By John D. Speth.

University of Chicago Press: 1983.

Pp.227. Hbk \$20, £17;

pbk \$8, £6.80.

OCCASIONALLY, a single study, by drawing together hitherto isolated areas of research, invites scientists in a discipline to take their work to a new level of synthesis and discourse. John Speth's *Bison Kills and Bone Counts* extends such an invitation to zooarchaeologists and to everyone who uses faunal evidence from archaeological sites to reconstruct past human diet. The book's subtitle, *Decision Making by Ancient Hunters*, signals its principal aim: to reconstruct the human decisions and actions that created an archaeological site, its internal structure and the patterning in its faunal remains. To this end, Speth applies the past decade's findings in vertebrate taphonomy and lithic replication studies to reconstruct the train of events that formed the Garnsey

site, a fifteenth-century bison-kill complex in New Mexico. He then draws upon ethnography, ethnoarchaeology and nutritional research to build a case for nutritionally conditioned selectivity on the part of the bison hunters in their butchery and use of male rather than female bison.

The distribution of bones and lithics along a small tributary of the Pecos River reflects successive kill and butchery episodes, spanning nearly a century. These kills, according to Michael Wilson's analysis of the bison's dentitions, took place in the spring. From sex ratios of crania, Speth suggests that the hunters focused on "bull groups", the usual social grouping of male bison except during rut. He explains the presence of females in bone clusters supposedly representing bull-group kills by noting that non-calving females join such associations among modern bison, although a kill of at least one cow-calf group, also in the spring, is implied by the presence of some calf remains.

Speth applies recent findings from vertebrate taphonomy in analysing the patterns of bone representation, state of preservation and damage patterns of the bison remains, inferring that the assemblage has been minimally affected by carnivores or

fluvial action. He concludes that patterns of selective bone representation at Garnsey are best understood as direct products of human processing activities.

Applying metrical discriminators of sex to both cranial and postcranial elements, Speth discerns striking differences in the representation of the two sexes' bones. While the ratio of male to female crania is 6:4 for the site as a whole, that of limb elements is reversed, 7:3 in favour of females. The ratio of the less portable and less useful crania probably reflects that of the animals killed, Speth contends, whereas the appendicular element ratio reflects the highly purposeful selection, and removal for eating, of the limbs of male animals, with discard of those of females.

To understand why these and other female elements, normally of moderate to high nutritional yield, were left at the site unused, Speth invokes sexually divergent regimes of fat storage and metabolism among ungulates. The anomalously over-represented female bones are in fact those most likely to be low in marrow fat and fatty acids under nutritional stress. Calving and lactating bison cows are more likely to be heavily stressed during the spring season than are males. Speth argues that the low fat and fatty acid content of such female limb segments would have reduced their attractiveness to human hunters, themselves at the most fat-depleted point in their own nutritional cycle.

To support his contention, Speth presents data on human nutrition in circumstances of seasonally low carbohydrate and fat, yet moderate protein availability. These data imply that fat, carbohydrates or both are so fundamental to the body's use of ingested protein and its own stored energy that, in such circumstances, human hunting and foraging strategies will emphasize their recovery. Since such seasonal scarcities characterize the lives of foragers in temperate to arctic regions, these implications should be of interest to anyone practising archaeology in these zones.

Bison Kills and Bone Counts is lucidly written. Its clarity, brevity and low cost in paperback recommend it particularly as a text in university courses on human diet or zooarchaeology. But any archaeologist should find it good reading, as should many students of human nutrition. □

Diane Gifford-Gonzalez is in the Department of Anthropology at the University of California, Santa Cruz.

● Abraham Pais's well-received biography of Albert Einstein, *'Subtle is the Lord . . .'*, has been issued in paperback by Oxford University Press. The book was reviewed in *Nature* (300, 116; 1982) by Arthur I. Miller who said of it: "The level of scientific presentation is high, but so are the rewards, and the book fills in one part of the mosaic that is Einstein, his life, his times and his science". Price of the paper edition is \$12.95, £6.95.

In between computing and psychology

Margaret A. Boden

Machine Learning: An Artificial Intelligence Approach.

Edited by Ryszard S. Michalski, Jaime G. Carbonell and Tom M. Mitchell.

William Kaufmann (distributed by W.H. Freeman): 1983.

Pp.572. \$39.50, £36.95.

The Psychology of Human-Computer Interaction.

By S.K. Card, T.P. Moran and A. Newell.

Lawrence Erlbaum: 1983. Pp.469.
\$39.95, £27.50.

ONE of the most intractable problems in artificial intelligence is machine learning. Admittedly, programs were written a decade ago which could "learn" (Winston's "concept-learner" is one of the best known). But all — and usually only — the relevant parameters had to be specifically predefined for these programs. Moreover, all the input samples (which were presented in a non-arbitrary order) were either clear examples or clear counterexamples of the concept concerned. For the human being, whether adult or baby, life is not so kind. Yet we are able to learn from the potentially confusing array of rich environmental input (of which what other people explicitly tell us is a very small part indeed).

If truly powerful machine intelligence is to be developed, similar feats of learning must be achieved. Quite apart from its theoretical interest, machine-learning is necessary from the point of view of practical applications. For we can control the input of information to computer-systems only to a limited extent, and what we know of human and animal learning suggests that countless presentations of slightly varying input are often needed for a system to learn.

The first of the two books reviewed here, *Machine Learning*, is the best source of discussions on these problems outside specialist conference-proceedings, and is indispensable for the serious student of the field. Its 16 chapters describe a variety of specific learning-systems in detail. Some of these are related to specific applications (such as commerce, experimental genetics and education), and most explore a particular theoretical approach to learning which, if valid, could be widely applied. For instance, an idea currently attracting much interest is the notion of "version spaces". This programming methodology is an attempt to overcome the limitations mentioned above, for it allows learning to progress by the gradual clarification of a "grey area" between the *known examples* and the *known non-examples*. Version-spaces (a concept largely developed by one

of the editors, T. Mitchell) are well described in this volume, as is D. Lenat's work on heuristics for discovery, and R. Michalski's model of inductive learning.

Several chapters are of a more general nature, focusing on underlying principles and providing comparative evaluation of different approaches. These overviews of the field are stimulating and valuable. There is also a useful 40-page classified bibliography of machine learning. The book even includes an impish salute to the Hegelian dialectic, in the form of a chapter by H. Simon which proposes "a thesis to which perhaps the other chapters serve as antitheses" — namely, the suggestion that the best way to approach artificial learning-systems is to concentrate on *human* learning and its simulation.

The second book under review is co-written by Simon's longstanding collaborator A. Newell, and addresses one of the five issues identified by Simon as "priorities" for research in machine learning. *The Psychology of Human-Computer Interaction* is an attempt to understand the performance of people using interactive computing systems.

While neither of these books is for the uninitiated, this one is the less accessible of the two. It presents extremely detailed performance-studies of specific interactive skills, based on the (production-system) approach to psychology which Newell and Simon have pioneered over the past 30

ADVERTISEMENT

RESEARCH IN BRITISH UNIVERSITIES POLYTECHNICS AND COLLEGES

RBUPC

1983

YOUR GUIDE TO
CURRENT RESEARCH IN THE UK

Vol. 1 Physical Sciences £44
ISBN 0 7123 2009 1

Vol. 2 Biological Sciences £44
ISBN 0 7123 2010 5

Vol. 3 Social Sciences £33
ISBN 0 7123 2011 3

Set of 3 volumes (1, 2 & 3) £109
ISBN 0 7123 2008 3



Available from:

Publications Section (B3),
British Library Lending Division,
Boston Spa, Wetherby,
West Yorkshire, LS23 7BQ

Circle No.29 on Reader Service Card.

years. For instance, it offers a model (simulation) of text-processing which predicts not only the precise time taken for a user to perform a particular task, but also the individual keystrokes involved. These concrete case-studies are used as the basis for drawing more general conclusions regarding psychological matters, and how they should be taken into account by computer scientists when designing artificial information-processing systems.

The authors are unashamedly interdisciplinary in their approach, saying "We see this book not just as a book in applied psychology, but as a book in computer science as well". They regard certain central aspects of computers as being as much a function of the nature of human beings as of the nature of the computers themselves. This is partly because it is human beings who have designed computers and who wish to use them. But it is also due to the fact that both human

minds and computers are information-processing systems, and as such conform to general principles of information-processing which (the authors claim) their approach promises to clarify.

Many things taken for granted in today's technology are not merely inefficient at the level of superficial ergonomics, but radically mismatch the basic organizing principles of mind and machine. This mismatch, according to Card *et al.*, is neither theoretically nor technologically necessary. For example, a more precise model of user-performance would have encouraged computer-engineers to design time-sharing systems in a significantly different way. It follows that (precise) models of the human user have a place in "what every computer scientist should know to call himself a computer scientist". □

Margaret A. Boden is Professor of Philosophy and Psychology at the University of Sussex.

Adding colour

R.S. Davidson

The Colour Science of Dyes and Pigments.

By K. McLaren.

Adam Hilger, Bristol/Heyden, Philadelphia: 1983. Pp.185.

£17.50, \$34.

Organic Chemistry in Colour.

By P.F. Gordon and P. Gregory.

Springer-Verlag: 1983. Pp.322.

DM178, \$66.

THE study of colour is one which bridges art and science, and industry and academe. In both of these books it is clearly evident that the authors are fascinated by the subject and truly knowledgeable. This, no doubt, in part reflects the fact that all three of them have spent many years working in the dyestuffs industry.

In his first chapter, McLaren paints a most vivid picture of the historical development of colouring matters, and indeed throughout the book he discusses modern achievements and current ideas in relation to concepts developed years ago. Overall, he is mainly concerned with colour science, that is, the definition of colour and the especially complex topic of its measurement.

To lead up to a description of the basis of modern colour measurement, McLaren covers the process of colour vision and the factors which influence perception of colour. Succeeding chapters deal with measurement of colour, colour and measuring instruments, quantification of colour difference and computer match prediction. Two other chapters, covering the photophysics and photochemistry of colouring matters, and the application of colouring matters, are fairly well integrated with the rest of the book.

McLaren writes authoritatively, yet in an easy, readable style, and entertainingly intersperses his text with anecdotes. For example, the ability of men who have taken up holy orders to devise palatable and health restoring drinks is well-known, but to find one who, as early as 1611, dreamt up a colour system is rather surprising (that man was Sigrid Aron Forsius, a Swedish monk and astrologer). The book is thoroughly up to date and can be highly recommended to students and practitioners of colour measurement. My only criticism, and at that a minor one, is that the structural formulae are badly drawn and in some cases are incorrect — perhaps the author doesn't like the organic chemistry aspects of his subject.

For that sort of information, one should turn to Gordon and Gregory, who deal in depth with the chemistry (synthesis and properties) of many classes of dye. In contrast to McLaren, they hardly mention colour measurement (even though a colour map is included in the book!). In many respects, therefore, the two books are complementary.

Gordon and Gregory also include an excellent chapter on the application and fastness of dyes. Another attractive feature is provision of a bibliography at the end of each chapter which lists references under specific headings. I was impressed by this book; it is virtually free of typographical errors, and the quality of writing is generally good. While it is unfortunate that the title is so ambiguous, and that the authors have not exploited the usefulness of dyestuff chemistry in teaching the principles of electrophilic and nucleophilic aromatic substitution reactions more fully, both teachers and students of colour chemistry will find it to be well worth reading. □

R.S. Davidson is Professor of Organic Chemistry at the City University, London.

Sinaean seaweeds

Ralph A. Lewin

Common Seaweeds of China.

Edited by C.K. Tseng.

Science Press, Beijing: 1983. Pp.316.

Price not known.

ALTHOUGH almost every country has one or more beautifully illustrated publications dealing with native flowers and native birds, few have comparable works with good colour photographs of the local seaweeds. Japan and Australia have such books: the United States and Switzerland (perhaps for different reasons) do not. Now the Chinese have published a lovely volume, in English, lavishly illustrated on heavy paper with 149 coloured plates showing 512 species, including most of the major algae of the coasts of China. (It is noteworthy that 14 of the plates deal with cyanophytes, which tend to be ignored altogether in most seaweed floras. Bacteriologists will doubtless be pleased.)

Common Seaweeds of China is not merely a coffee-table production, however. For every species included, some 50–60% of the total recorded in the Chinese marine algal flora, there is a colour photograph and a 5–10-line description, in technical but clear scientific style, with details of form and structure, and plenty of dimensions of cells and organs (so necessary for identifications, yet all too often sadly lacking in other algal texts). There is information on distribution on the Chinese coasts and elsewhere, and, where relevant, mention is made of culinary or other uses. The generic and specific binomials are accompanied by full authorities. The English syntax and spelling are almost impeccable. A bibliography with 146 references provides a most valuable introduction to the Chinese seaweed literature for phycologists elsewhere in the world.

In his preface, C.K. Tseng explains that the book is not intended to be a working manual. There are few technical details whereby genera of seaweeds are to be distinguished and no diagnostic keys. It could hardly be otherwise in a book of this sort. Tseng also solicits comments and suggestions for future editions, but it is hard to find fault except with some of the photographs which are really very poor. I recognize that many of the species are not easy to photograph clearly, especially the filamentous forms, but I think it should be possible to do better.

Common Seaweeds of China is altogether quite an achievement. Tseng and his colleagues have produced a remarkable book, one for which all English-speaking phycologists will be grateful. □

Ralph A. Lewin is a Professor at the Scripps Institution of Oceanography, University of California at San Diego.

Pittsburgh Conference

The companies featured in this week's pages are all exhibiting at the Pittsburgh Conference, to be held in Atlantic City, New Jersey, USA from 5 to 9 March 1984

● Graphic Information Systems is exhibiting the latest in its range of Imagan image analysis systems at the Pittsburgh Conference. Quantitative analyses of medical conditions can be made from photographs and slides. The new Imagan system provides a range of statistical and stereological analysis packages.

Circle No. 100 on Reader Service Card.

● A 5'-labelling kit is available from Boehringer Mannheim in which the reaction conditions have been optimized to give maximal labelling of the 5'-hydroxyl terminals of DNA. In the direct phosphorylation reaction, the amount of DNA and enzyme has been adjusted to maximize dephosphorylation and phosphorylation. Using the recommended ratio of ATP to 5'-hydroxyl ends of 5:1, the majority of the DNA molecules receive a ^{32}P -group at their 5' ends. Each kit allows up to 200 μmol of the control DNA fragments to be labelled. Complete working instructions are available on request.

Circle No. 101 on Reader Service Card.

● Hach has introduced a new combination pH electrode. The new model is designed to reduce the problems associated with combination electrodes operating at low temperatures caused by a build-up of solids in the porous plug. Hach's injection flow combination electrode reduces this and can operate at sub-zero temperatures.

Circle No. 102 on Reader Service Card.

● Schoeffel McPherson's model 251 spectrometer, which is being exhibited at Pittsburgh, is available in the UK through Intersci.

Circle No. 103 on Reader Service Card.

● Labconco has designed a stainless steel hood for use when handling perchloric acid. The perchloric acid hood has an integral washdown system which helps prevent accumulation of perchlorates. The seamless interior has no screws or hollow fasteners to contain any residue. The hood has an air foil and automatic air bypass to assure continued sweeping action and to give sufficient velocity for fume removal.

Circle No. 104 on Reader Service Card.

● The first in a new line of monoclonal antibodies has been introduced by Boehringer Mannheim. The type I antibody is specific for choline acetyltransferase and may be used for the immunohistochemical localization of choline acetyltransferase by indirect immunofluorescence, avidin-biotin or peroxidase-antiperoxidase (PAP) methods. Cross reactivity is shown with choline acetyltransferase from humans, rats, chickens and pigs. The preparation is available in a stabilized, lyophilized form.

Circle No. 105 on Reader Service Card.

● The new SDS calibration protein set from Boehringer Mannheim covers a molecular weight range of 20,000 to 340,000. The five proteins are individually packaged as solutions in glycerol (0.4 mg each). Up to 80 runs may be performed on most of the usual gel systems (Shapiro, Weber and Laemmli) with this kit.

Circle No. 106 on Reader Service Card.

● The new Varian 3400 gas chromatograph (GC) is small and easy to use. It has a built-in microprocessor that controls all the GC functions including autosampler control, external event relays, of which there are four, multi-ramp oven temperature programming, detector range and auto-zero. The VA 3400 has bidirectional RS-232C communication and can be mated with the Vista 402 data system.

Circle No. 107 on Reader Service Card.

● Two new glycohydrolases have been introduced by Boehringer Mannheim for use in the removal of sugars from glycoproteins and in the analysis and structural study of glycopeptides. Endoglycosidase D hydrolyses glycopeptides belonging to the complex chain group, such as unit B glycopeptides of thyroglobulin and glycopeptides from cell membranes. Endoglycosidase H hydrolyses core dinitroacetylchitobiose linkages in high-mannose glycopeptides such as ovalbumin glycopeptides.

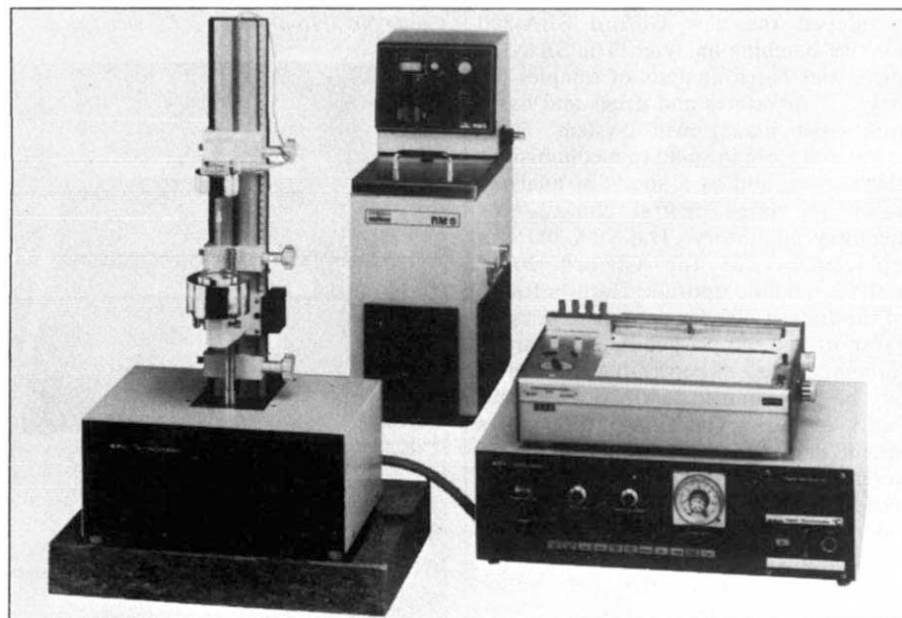
Circle No. 108 on Reader Service Card.

● Methylase *HpaII* which is free from *HpaI* methylase, *HpaI* restriction endonuclease, *HpaII* restriction endonuclease and non-specific DNases can now be obtained from Boehringer Mannheim. The preparation is supplied in Tris buffer at a volume activity of approximately 1 unit per μl .

Circle No. 109 on Reader Service Card.

● Brinkmann Instruments has introduced the Lauda Tensiometer. This automatically measures and records the surface tension of liquids and the interfacial tension between two liquids as a function of time. The Tensiometer consists of a measuring device and control unit (power supply/amplifier). The de Nouy ring and Wilhelmy plate are used as measuring probes. A recorder and a constant temperature circulator are necessary accessories.

Circle No. 110 on Reader Service Card.



Brinkmann's Instruments' Lauda Tensiometer

These notes are based on information provided by the manufacturers. For further details circle the appropriate numbers on the Reader Service Card bound inside the journal.

● The Wenking Potentio-Galvano-Scan model PGS 81, a combination medium power potentiostat (galvanostat)/high-accuracy scanner, is now available from Brinkmann. The model PGS 81 records current-density against potential characteristics. Potential and current settings can be selected in the range of $\pm 5V$ ($\pm 1A$). Buffer-converted and grounded low impedance outputs are provided for connection to x-y recorders or oscilloscopes.

Circle No. 111 on Reader Service Card.

● J.T. Baker Research Products can now supply the Bakerbond chiral phase (DNBPG) column. This consists of high enantiomeric purity (*R*), *N*-3,5-dinitrobenzoyl-phenylglycine that is bonded covalently to 5- μm aminopropyl silica gel. Due to the covalent bonding of the chiral stationary phase, the column may be used with high-polarity mobile phases.

Circle No. 112 on Reader Service Card.

● Elsevier is introducing a series of software packages dealing with analytical chemistry. Four programs are being introduced initially. These are Refvalue, Balance, Cleopatra and Instrumentune-up. Refvalue is a software package that calculates reference intervals from total hospital patients laboratory data, and Balance is a program to compare two series of measurements. This package includes classical parametric and non-parametric tests. These can be used to investigate whether, for instance, two medical treatments, two drugs, or two analytical methods yield the same result. Cleopatra is an extendable set of programs covering the chemometrics library which is intended to be an aid in teaching and research. Instrumentune-up is an experimental optimization program.

Circle No. 113 on Reader Service Card.

● Corning Medical and Scientific has introduced the new Gilford SBA-300 selective batching analyser. The SBA-300 carries out batch analysis of samples for enzymes, substrates and drugs and has a work load management system. It is intended for use in small to medium-sized laboratories, and as a specialist analyser within the large central clinical biochemistry laboratory. The SBA-300 is a replacement for the Gilford 203-S analyser, and incorporates features found on the Impact 400 for data management, transport and sample handling. A variety of method types and operating parameters may be programmed, and stored permanently on disk. The floppy disk stores patient data and test results (up to 255 patients with 6,000 results) with the information being automatically collated at the end of each analysis. The SBA-300 has throughput rates of around 100 tests per hour for enzymes and 250 tests per hour for end-point chemistries.

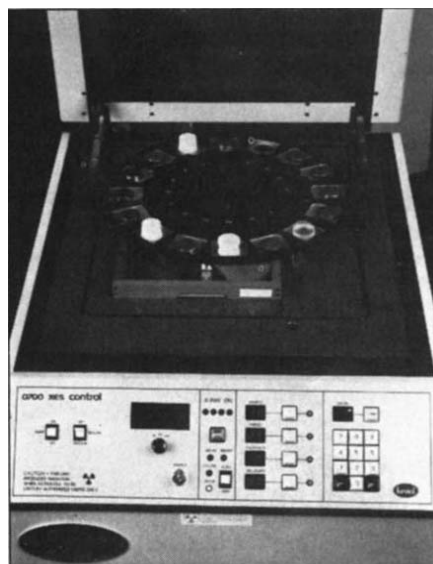
Circle No. 114 on Reader Service Card.

● Bio-Rad now offers ion exchange chromatography columns packed with a new medium designed for the analysis and purification of ionizable biochemical substances. This new medium is a derivative of the Bio-Sil TSK SW packing material and provides a complement to the Bio-Sil TSK SW gel filtration columns. The ion exchange matrices are hydrophilic and exhibit low hydrophobic adsorption. The rigidity of the medium allows rapid regeneration at high flow rates. Aqueous buffers and most organic solvents in the pH range 2.5–7.5 can be used without shrinking and swelling. Denaturants and non-ionic surfactants can also be used.

Circle No. 115 on Reader Service Card.

● The 0700 laboratory X-ray fluorescence system from Kevex provides a quantitative method of simultaneous multi-element analysis that does not alter the chemical or physical properties of the sample. Up to 24 samples can be loaded into the 0700 and analysed using direct, filtered direct or secondary excitation. All the 0700's accessible functions are under microprocessor control and manual control consists of direct instruction to the microprocessor via the 0700 front panel keypad. User and microprocessor selectable parameters include: sample position, direct or secondary target X-ray tube position; X-ray high voltage, X-ray tube current and transmission filter/collimator selection. When low mass samples need to be analysed, such as filter papers or thin film substrates, a large collimator can be placed between the tube and the sample. As both systems ensure saturation of the detector/counting electronics, there is no reduction in the counting ratios or stabilities. A set of six pure element targets is provided for secondary excitation and these can also be used in conjunction with filters to remove residual scattered tube continuum. The system has a combined mechanical and counting error of 0.5%.

Circle No. 116 on Reader Service Card.



The 0700 system from Kevex

● Perkin-Elmer's chromatography laboratory automation software (C/LAS) package is designed to give chromatography automation. The C/LAS module is the newest addition to Perkin-Elmer's laboratory information management system (LIMS/2000). It can be used either as a stand-alone system for automated chromatography runs or can operate as an integral part of LIMS/2000 with complete links to the LIMS/2000 database. Using C/LAS, data can be collected from up to 72 chromatographs and passed through an analog-to-digital converter to a Perkin-Elmer 3200 series minicomputer for data reduction. The system features a 32-bit minicomputer and a buffered interface for high-speed data collection on up to 72 channels, interactive colour graphics and a range of terminals, graphics workstations, printers and plotters.

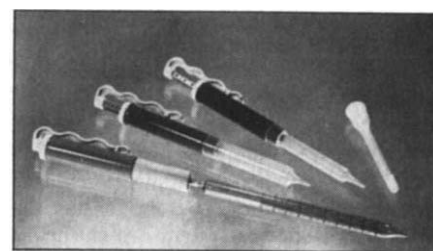
Circle No. 117 on Reader Service Card.

● The TIME (twelve-way inoculating made easy) dispenser from Dynatech Laboratories is a sterile, disposable polystyrene dispenser which delivers 100 μl per well of liquid into a 96-well Microtiter plate. The dispenser may be used in feeding cell cultures, dispensing aetiological agents and delivering conjugates in ELISA methods. With each stroke of the plunger, 12 wells are filled at once with 100 μl of liquid. The built-in reservoir holds enough liquid to fill a 96-well plate.

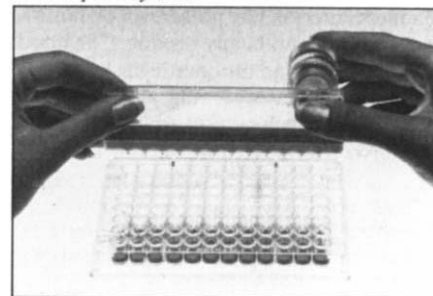
Circle No. 118 on Reader Service Card.

● The Sarpette pipetting sample handling system is now available from Sarstedt. Sarpettes are colour-coded and provided in two models, both of which can be used with disposable Sarpette tips or adapted for standard glass or plastic pipettes. Model I Sarpettes are available in three sizes whilst model II Sarpettes come in six sizes and have click stop calibrations for repetitive proportional dispensing of volumes in identical aliquots.

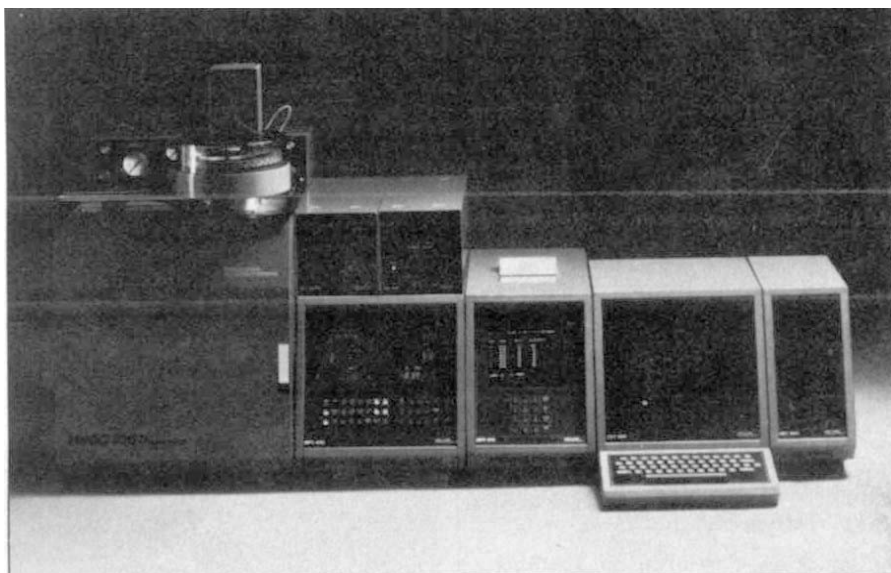
Circle No. 119 on Reader Service Card.



The Sarpette system



Dynatech's TIME dispenser



The Mega series gas chromatograph from Erba Instruments

● Erba Instruments is introducing a new Mega series gas chromatograph (GC) for high resolution chromatography at the Pittsburgh Conference. The Mega series has a new oven design which, when combined with new columns, allows for a 'thermally quiet' zone where temperature readings can be carried out to within 0.05% of actual temperature. The system can be supplied with a microprocessor-based control system or a fully-programmable computer. Features of the series include injector/detector systems for packed columns and Grob split/splitless and cold on-column capillary injectors. The system can be adapted for automatic operation by adding the automatic liquid injector (AS-V 570) or automatic headspace analyser (HS 250).

Circle No. 120 on Reader Service Card.

● Spectra-Physics is showing its line of liquid chromatographs, gas chromatographs, and data systems at the Pittsburgh Conference. The company will also demonstrate a new Labnet computer which enables a chemist to control an entire chromatography laboratory from a single work station. The data systems are supplied in customized forms using the new Spectra-Physics special purpose software programs, the (SP)³ series of pre-programmed chips. The (SP)³ chips can be plugged into the instrument, to customize it for applications such as automated separation optimization (Optim I), size exclusion chromatography (GPC), and UV/visible spectra acquisition (Autoscan). The chips can be used with the SP8100 automated HPLC system, the SP8700 modular ternary HPLC system and the SP8770 high performance isocratic liquid chromatograph. Also available in the (SP)³ series are the autosampler control program and boiling range distribution (SIMDIS) program for the SP7100 gas chromatograph.

Circle No. 121 on Reader Service Card.

● Nelson Analytical's new model 6000 data system joins the company's dedicated chromatography software to the new HP 1000 series Micro 26 minicomputer. The model 6000 can handle more than 30 chromatographs simultaneously, each generating up to 100 data points per second. It is compatible with Nelson Analytical's chromatography data systems, which are based on HP 200 series desktop computers. The system can run 256 program tasks simultaneously. The standard model 6000 system consists of a processor, a graphics terminal, Nelson Analytical instrument interfaces and the model 6000 software.

Circle No. 122 on Reader Service Card.

● An acrylamide medium is now available from Gallard Schlesinger. The product has an acrylic acid content of less than 0.001% and contains less than 0.001% of insolubles. This electrophoresis grade reagent allows for production of clear solutions and transparent gels and is available in 100 and 500 g packages.

Circle No. 123 on Reader Service Card.

● The CALS (computer automated laboratory system) from Beckman can control laboratory management, data handling, reports, storage and retrieval of data. CALS acquires data from chromatographs, spectrometers and physical test equipment, accepts analog or digital outputs and will take in manually-entered data via cathode ray tube terminals. The instrument's functions include sample login, automatic instrument operations and data acquisition; data display; multi-point calibration and comparison with standards. The system can be expanded. The software options offered include auto-analyser data analysis, calibration and reporting, sample tracking, status monitoring and laboratory data base management.

Circle No. 124 on Reader Service Card.

● Rapid chemical or biochemical reactions can be measured by UV/visible absorption or fluorescence using the Dionex D-100 series stopped-flow spectrophotometers and associated accessories. The instruments are based on the Gibson design and are used to measure reaction time constants using small sample sizes. They operate in the range of milliseconds to minutes. The system features optics for a wide range of wavelengths and a Kel-F flow system for maximum chemical inertness.

Circle No. 125 on Reader Service Card.

● A centre well flask is available from Kontes that can be used for metabolic studies of biological materials. The unit consists of disposable polypropylene centre wells and double-seal rubber stoppers, and reusable 10 or 25 ml borosilicate Erlenmeyer flasks. The stoppers will accommodate any vessel with a 14 mm mouth. Applications include measuring C¹⁴O₂ yields following incubation, *in vitro* incubations, insulin bioassays, persulphate oxidations and the distillation of any volatile acid or base.

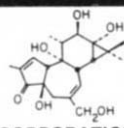
Circle No. 126 on Reader Service Card.

● The Response UV/visible scanning spectrophotometer is now available from Gilford. Using a 16-bit microprocessor, the instrument's autoranging photometric system provides up to 40 readings per second. The standard software package includes wavelength scanning, kinetics, gel scanning, multiwavelength scanning and time scanning routines. Accessories are available that can accommodate 1, 5 and 10 cm cuvettes, rapid sampling and flow cell accessories and gels of up to 20 cm.

Circle No. 127 on Reader Service Card.

ADVERTISEMENTS

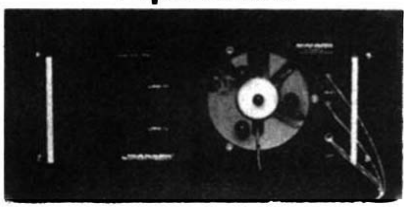
PHORBOL ESTERS



LC SERVICES CORPORATION
165 New Boston Street
Woburn, MA 01801
(617) 938-1700

Circle No. 32 on Reader Service Card.

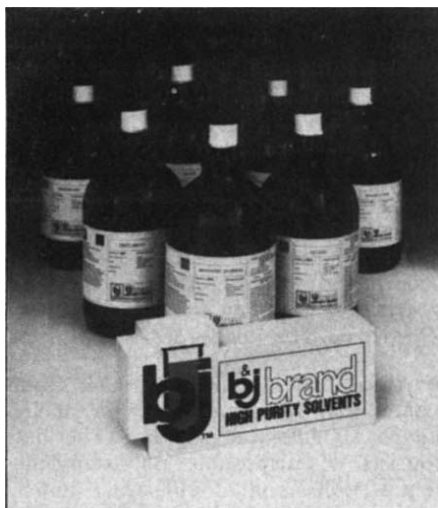
Liposomes



LIPOPREP the equipment for the reproducible preparation of stable unilamellar liposomes of defined size in the range of 30-300 nm ϕ with extremely narrow particle size distribution for e.g. the efficient incorporation of (therapeutic) agents incl. antibodies.

DIANORM-Geräte, P.O. Box 126, D-8 München 65, FRG

Circle No. 35 on Reader Service Card.

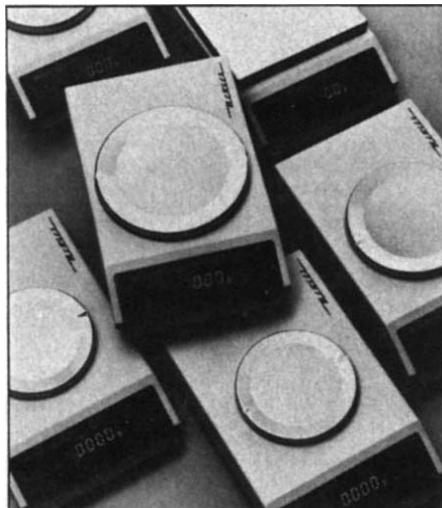


Burdick & Jackson's solvents

● Burdick & Jackson Laboratories has added decahydronaphthalene and 1,2,4-trichlorobenzene to its line of B&J brand solvents for gel permeating chromatography (GPC) and polymer additive analysis. For high temperature gel permeating chromatography (HTGPC), trichlorobenzene is an alternative to *o*-dichlorobenzene for the analysis of polyolefins and other polymers. Decahydronaphthalene is useful primarily for determination of polymer additives such as BHT and also for GPC of polymers. *Circle No. 128 on Reader Service Card.*

● The Magiscan 2 is a new image analysis system from Joyce-Loebl which is driven mostly by software. It extracts and analyses information not only from optical or electron microscopes but also from photographs, negatives, film or real life images. The system is now available with pseudo-colour. The image processing itself is performed by selecting options from an interactive menu which guides the user through the correct sequence of commands. During routine analysis, the operator may change the illumination or other conditions in order to achieve accurate results. These results may then be used in graphic or standard statistical form on the television screen or printer. *Circle No. 129 on Reader Service Card.*

● The Astramax ultrasonic cleaning and vapour degreasing systems are now available from Heat Systems Ultrasonic. These units use ultrasonic cavitation to dislodge contaminants. The Astramax is available in several sizes with either 25 kHz or 40 kHz generators. Separate generator and transducer systems are available to add ultrasonic energy to existing tank systems. Cleaning compounds for use with the system are also available from the manufacturer. Ultrasonic vapour degreasers are designed to be used with volatile solvents such as freon or trichlorethylene when the contaminant is not water soluble or when rapid drying of the part is required. *Circle No. 130 on Reader Service Card.*

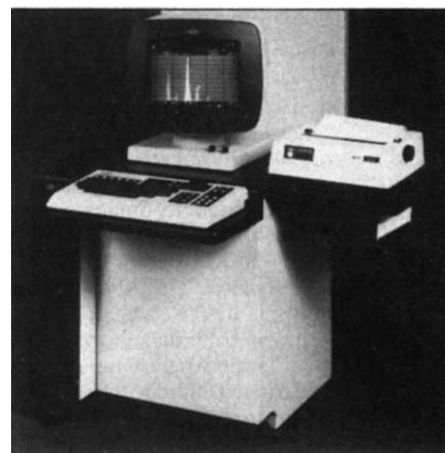


Mettler's balance

● Mettler has developed the LabPac, a system that converts standard laboratory balances into multifunction balances. The basic instrument is a Mettler PE balance although the actual LabPac consists of a GE305 application input device and four program keys for determining net total, percent determinations and percentage weighings, animal weighing, mean value ($\bar{x}$) and standard deviations and result indications in non-metric weight units. Depending on the key that is inserted, the balance is programmed for the desired application. Each balance is operated with the time-tested Mettler automatic single control bar and the three buttons of the GE305 application input device. The unidirectional data output (current loop and RE232C data interface standards) allows data to pass to peripheral instruments. *Circle No. 131 on Reader Service Card.*

● Pharmacia Fine Chemicals has developed a new pH monitor which, when used in conjunction with its flow-through pH electrode, gives continuous pH monitoring in liquid chromatography. The pH monitor with a flow cell is suitable for both fast protein liquid chromatography (FPLC) and HPLC. The dead volume of the flow-through pH electrode is 30 μ l which prevents the remixing of previously separated components and guarantees a reliable reproducible recording of pH changes in column eluent. The monitor's liquid crystal display shows the correct pH at all times and also plots the pH continuously over a chosen range. The incorporated span function enables the pH measuring range to be adjusted between 0.05 and 14 pH. The pH monitor also incorporates a high input impedance for amplifying and measuring small signals which is accurate to within ± 0.01 pH units. Fluctuations in temperature give a low pH drift of < 0.005 pH units per $^{\circ}\text{C}$. The flow-through pH electrode consists of a flow-cell and a fast-response, epoxy-bodied microelectrode screwed together to form a single unit. *Circle No. 132 on Reader Service Card.*

● Kratos Analytical Instruments has introduced a new development in mass spectrometry: the MS 50 TC. The spectrometer has a 7,500 mass range at full power (due to new inhomogeneous field optics), and a turbomolecular pump. Other important features include a range of inlet systems such as gas chromatography, liquid chromatography, post acceleration detector for high mass FAB and self-diagnostic facilities. The system also has touch control switches, a central microprocessor for setting and controlling all mass spectrometry functions and a keyboard controlled autoconsole. *Circle No. 133 on Reader Service Card.*



The Kevex Microanalyst 8000

● The Kevex Microanalyst 8000 is a full-signal qualitative and quantitative microanalysis system. It produces line profiles and digital images using X rays, back-scattered electrons, transmitted electrons and Auger electrons. It can also accept wavelength dispersive X-ray signals as well as energy dispersive signals. The Microanalyst 8000's analytical power is derived from the Quantex-Ray II software, which can control automatic peak detection and identification routines, a choice of background-fitting programs, peak-fitting routines for complex deconvolutions and can carry out Magic V ZAF corrections using reference standards. In order to speed up data analysis, the Quantex-Ray II package can be used with Hyperdrive. This is a solid-state mass storage system which allows memory access to be 25 times as fast as conventional disk storage systems and offers a 4-Mbyte storage capacity. Its transfer speed is 380 kbyte s^{-1} and there is no latency. The conventional disk system is preserved for lower priority or large volume storage. The Microanalyst 8000 has a colour display with a matrix of 512 x 256 individually-addressable pixels. The instrument can be controlled vocally. The Microanalyst 8000 has a unique voice recognition capability which allows it to recognise and execute spoken commands. All the operator has to do is to teach the computer the vocal instruction it must recognise. *Circle No. 134 on Reader Service Card.*

Awards

The winners of the Feldberg Foundation awards for 1984 are **Professor E. Muscholl** of the Department of Pharmacology, University of Mainz, FRG, and **Professor D. Weatherall** of the Nuffield Department of Clinical Medicine and honorary director of the MRC Molecular Haematology Unit, Oxford, UK.

The Association for Radiation Research has awarded the Weiss medal jointly to **Professor D. Schulte-Frohlinde** and **Professor C. von Sonntag**, both of the Institut für Strahlenchemie im Max-Planck Institut für Kohlenforschung, Mulheim a.d. Ruhr, FRG, for their work on the radiation chemistry of DNA and its components.

Nominations are invited for the King Faisal International Prizes in medicine and science. The prize for medicine is to be awarded for work on viral hepatitis and the prize for science is for contributions to biochemistry. Each prize consists of a certificate, a medal and the sum of 250,000 Saudi Riyals (£51,000). Nominees must satisfy the following conditions: they must have published an outstanding academic work in the relevant subject which will lead to the benefit of mankind and the enrichment of human thought and they must be leading members of recognised educational institutions of world-wide fame. Works submitted for the nomination must not have been awarded a prize by any international educational establishment. Ten copies of nominations, which must contain details of the nominee's entire academic history since this will be taken into account when awarding the prize, must be submitted by 9 August, 1984. Further information is available from the Secretary General of the King Faisal International Prize, PO Box 22476, Riyadh 11495, Kingdom of Saudi Arabia.

Meetings

12–13 April, **Meeting on Determinants of Phytoplankton Distribution**, Swansea, UK (Professor P.J. Syrett, Department of Botany and Microbiology, University College of Swansea, Singleton Park, Swansea SA2 8PP, UK).

25 April, **The William Smith Meeting of the Geological Society of London — A Theoretical Framework for the Pleistocene Ice Ages**, London (Ms Caroline Symonds, Administrative Secretary, Geological Society of London, Burlington House, London W1V 0JU, UK).

8–9 May, **European Meeting of the Society for Environmental Geochemistry and Health**, Birmingham, UK (Dr B.E. Davies, Department of Geography, University College of Wales, Penglais, Aberystwyth, Dyfed SY23 3DB, UK).

20–24 May, **Annual Conference of the American Society of Photographic Scientists — State of the Art in Image Technology**, Boston (Mr Robert Wood, 7003 Kilworth Lane, Springfield, Virginia 22151, USA).

21–23 May, **International Congress on New Compounds in Biological and Chemical Warfare: Toxicological Evaluation**, University of Ghent, Belgium (Professor A. Heyndrickx, Department of Toxicology, State University of Ghent, Hospitaalstraat 13, B-9000 Ghent, Belgium).

23–25 May, **International Symposium on Electric and Magnetic Separation and Filtration Technology**, Antwerp, Belgium (K.VIV, Symposium Electric and Magnetic Separation, Jan van Rijswijklaan 58, B-2018 Antwerp, Belgium).

1–3 June, **Symposium on Cancer and Genetics in 1984**, Essen, FRG (Professor Eberhard Passarge, Vice President of the European Society of Human Genetics, Department of Human Genetics, University of Essen, FRG).

3–8 June, **Conference on Water Chlorination: Environmental Impact and Health Effects**, Williamsburg, Virginia (Mr Robert Jolley, Oak Ridge National Laboratory, PO Box X, Oak Ridge, Tennessee 37830, USA).

3–8 June, **Copenhagen International Symposium on Osteoporosis**, Copenhagen (Mr Claus Christiansen, Department of Clinical Chemistry, Glostrup Hospital, University of Copenhagen, 2600 Glostrup, Denmark).

4–8 June, **School on Mathematical Modelling in Physiology and Clinical Medicine**, University of Padua, Italy (Professor C. Cobelli, Department of Electrical Engineering and Electronics, University of Padua, Via Gradenigo, 6A 35131 Padova, Italy).

7–10 June, **International Symposium on Rapid Methods and Automation in Microbiology and Immunology**, Berlin (Professor K.-O. Habermehl, Institute of Clinical and Experimental Virology of the Free University of Berlin, Hindenburgdamm 27, D-1000 Berlin 45, FRG).

9–15 June, **International Cystic Fibrosis Congress**, Brighton, UK (9th International CF Congress 1984, Cystic Fibrosis Research Trust, 5 Blyth Road, Bromley, Kent BR1 3RS, UK).

13–23 June, **International Wetland Conference**, Trebon, Czechoslovakia (Dr Dennis Whigham, Smithsonian Environmental Research Center, Box 28, Edgewater,

Maryland 21037, USA).

17–20 June, **Annual Meeting of the Society for Cutaneous Ultrastructure Research**, Helsinki University, Finland (Dr Kirsti Maria Niemi, Secretary of the Organizing Committee, Department of Dermatology, Helsinki University Hospital, Snellmaninkatu 14, Helsinki 17, Finland).

18–20 June, **International Conference on Chromatography and Mass Spectrometry in Biomedical Sciences**, Milan, Italy (Dr Alberto Frigerio, Italian Group for Mass Spectrometry in Biochemistry and Medicine, Via Eustachi 36-20129 Milan, Italy).

18–22 June, **International Symposium on Nuclear Power Plant Outage Experience**, Karlsruhe, FRG (International Atomic Energy Agency, Wagramerstrasse 5, PO Box 100, A-1400 Vienna, Austria).

19–22 June, **International Conference on Abundance Ratios in the Solar System**, Paris (Centre National d'Etudes Spatiales, Département des Affaires Universitaires, 18, avenue Edouard-Belin, 31055 Toulouse, France).

19–23 June, **Congress of the Collegium Internationale Neuropsychopharmacologicum**, Florence, Italy (Dr Giorgio Racagni, Institute of Pharmacology and Pharmacognosy, University of Milan, Via Andrea del Sarto, 21-20129 Milan, Italy).

21–23 June, **International Workshop on Voice Prosthesis**, University of Wuerzburg, FRG (Professor I.F. Hermann, Department of Otolaryngology, University of Wuerzburg, D-8700 Wuerzburg, FRG).

24–28 June, **Physiological and Pharmacological Control of Nervous System Development**, Chieti, Italy (Organizing Secretariat, Fondazione Giovanni Lorenzini, Via Montenapoleone, 23-20121 Milan, Italy).

24–29 June, **Course on Constructivism Today, Theory and Research**, Geneva (Archives Jean Piaget, Université de Genève, 6, rue de Saussure, 1211 Genève 4, Switzerland).

25–29 June **Conference on Chemical Research Applied to World Needs**, The Hague, The Netherlands (CHEMRAWN III Congress Bureau, QLT Convention Services, Keizersgracht 792, 1017 EC Amsterdam, The Netherlands).

26–30 June, **Symposium on Science in Egyptology**, Manchester, UK (The Manchester Mummy Research Team, Manchester Museum, The University, Manchester M13 9PL, UK).

27–30 June, **Symposium on Nucleosynthesis and Acceleration of Cosmic Rays**, Graz, Austria (Lydie Koch-Miramond, CEN Saclay, DPhG/Sap — Bâtiment 28, 91191 Gif sur Yvette, France).

Improving job prospects ahead?

Richard Pearson*

The long-awaited upturn in the UK job market could be in sight. But graduates may be faced with a limited choice.

THE latest forecasts for UK graduate job prospects in 1984 show a cautious growth in demand. The oil and chemical industries are showing some signs of improvement, while general engineering, manufacturing and civil engineering may be picking themselves up from the doldrums and increasing their intake of graduates. In the service sector there seems to be some further growth in demand for jobs in retailing, although banking, accountancy and central and local government, all major recruiters, seem to have stabilized their intakes. Higher education still has few jobs for new graduates and postgraduates; those that do appear are still predominantly short-term contracts.

It is the electronics and computer industries which continue to be the growth sectors, with some major UK employers doubling their number of graduate vacancies in 1984 when compared to the "flat" years of 1981-83. For example, one of the multinationals is increasing its planned intake from 200 to over 500. While electronics graduates are profiting the most from this upturn, there is a spill-over effect as employers widen their net to include physicists and mathematicians. Employers are also seeking small numbers of chemists and materials scientists. In computing, demand is still growing, and most of the vacancies are likely to be available to graduates from any discipline. It must not be forgotten, however, that some major companies were unable to find sufficient recruits in 1983.

The slight improvement in demand has to be set against the continuing increase in the number of new graduates seeking jobs. The number graduating with first degrees from the universities is expected to fall for the first time in 1984, down by perhaps 3 per cent to about 65,000. The polytechnics, however, still have a growing number of graduates, around 25,000 in 1984, to whom should be added over 20,000 postgraduates and another 10,000 from the other colleges of higher education, the Open University and so on. In total there could well be more than 120,000 students graduating in the United Kingdom in 1984, an increase of about 2 per cent over 1983. Within the total, about one in four will have a science degree, slightly more than engineering degrees. Some 10,000 graduates from 1983 and earlier will still be seeking first appointments.

Of course not all this year's graduates

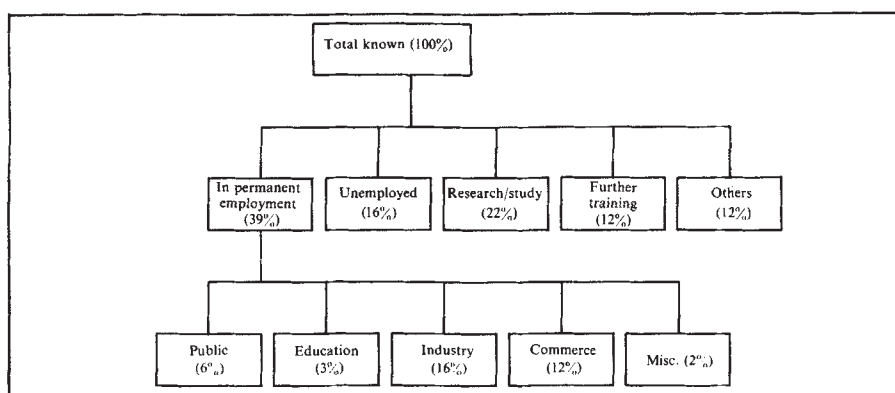


Fig. 1 Destinations of science graduates from UK universities in 1982.

will be seeking jobs. In recent years, one in three UK university graduates with a first degree in science has gone on to some form of further training or research, while another one in eight has returned overseas, or otherwise been unavailable for work. In total an average of about 40 per cent have gone directly into employment. Of the 1982 science graduates, most went into industry (10 per cent), the public service (7 per cent) and education (3 per cent) while commerce accounted for 10 per cent (Fig. 1). Only a few, however are likely to get jobs directly using their scientific knowledge (see *Nature* 306, 410; 1983). The polytechnic graduates show broadly similar destination patterns.

While a sixth of this year's science graduates may still be unemployed six months after graduation, a pattern echoed in many countries, they are still likely to have fared better than those leaving school or further education. In salary terms, however, over the past seven years they have slipped slightly behind the growth in average earnings. In 1975, the average basic starting salary for a new graduate was £2,475, which represented about 78 per cent of average earnings. In 1983 the figure was just over £6,000, an increase over the

period of over 140 per cent, but a decline to 69 per cent of average earnings (see Fig. 2). Until the mid-1970s, most graduates tended to receive broadly similar starting salaries, but since then market forces have come more strongly into play. Starting salaries now are often differentiated by subject, type of job, location, industry, and of course by the more personal factors. Not unnaturally, the biggest differentials tend to be between those industries where demand outstrips supply and those where it does not. In 1983, salaries ranged from under £5,000 to over £7,000, with those entering electronics-related jobs often getting well in excess of £7,500. Those going into jobs involving extensive further training often received nearer £5,000. Allowances for London working ranged from £1,000 to over £1,600, but some employers offered no such differential.

For postgraduates, a quick scan of *Nature's* job advertisements shows salary ranges of £5,500 to £8,000, with those lucky enough to get lecturing posts starting at just over £7,000, although these often require relevant experience as well as a higher degree. Outside the academic world, and research where postgraduate qualifications are normally regarded as essential, employers are usually far less interested in higher degrees and postgraduates have to compete in the graduate job market. While some employers may pay an increase on the basic starting salary of perhaps £200 for each additional year of study many give them no special financial recognition and in many jobs prefer a first-degree graduate.

The market reality for 1984 is, however, that UK graduates' main priority will continue to be that of finding any "suitable" employment. Only a minority will be able to select between interesting jobs with career prospects and to have any real say in their starting salary.

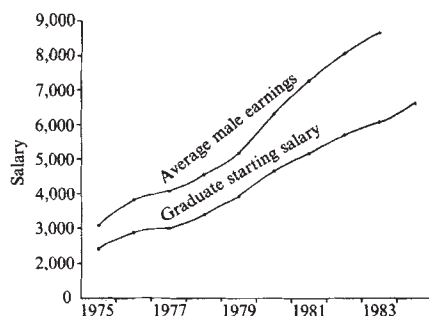


Fig. 2 Graduate starting salaries and average earnings in the United Kingdom from 1975 to 1984 (forecast for 1984).

*Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton BN1 9RF, UK.